The Journal of Applied Psychology

Vol. V JUNE, 1921 No. 2

EDITORS

JAMES P. PORTER
Clark University

WILLIAM F. BOOK
Indiana University

And a Board of Co-operating Editors

CLARK UNIVERSITY
WORCESTER MASS.
FLORENCE CHANDLER, Publisher

Entered as second class matter June 9, 1917, at the post office at Worcester, Mass. under the Act of March 3, 1879.

Acceptance for mailing at special rate of postage provided for in section 1103, Act of October 3, 1917, Authorized April 28, 1921.

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. V. Bingham, Carnegie Institute of Technology; Jesse E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; Joseph Jastrow, University of Wisconsin; Edmund S. Lanzfeld, Harvard University; James H. Leuba, Bryn Mawr College; Henry C. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, University of Wisconsin; Edward K. Strong, Jr., George Peabody College for Teachers; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; J. B. Wallace Wallin, Psycho-Educational Clinic, St. Louis; G. M. Whipple, University of Minnesota; Helen Thompson Woolley, Vocation Bureau, Cincinnati.

The scope of the JOURNAL OF APPLIED PSYCHOLOGY will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeal. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environmental conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or musical instruments, sports, etc.

VOLUME V JUNE, 1921 NUMBER 2

TABLE OF CONTENTS

H. T. Moore and A. R. Gilliland. The Measurement of Aggressiveness 97

Charles E. Bedaux. The Bedaux Unit Principle of Industrial Measurement 119

Elsie Oschrin Bergquist. A Study in Industrial Psychology—Tests for Special Abilities 127

L. Dewey Anderson. Estimating Intelligence by Means of Printed Photographs 152

W. F. Book. Two Cases Showing Marked Change in I. Q. 156

G. W. Fiske. A Preliminary Study of the Correlations Between Classification of Political Traits and the Results from the Downey Will-Profile 159

Lewis M. Terman and Jessie C. Fleming. Preliminary Report on a Gifted Juvenile Author 163

Abe Hart Arlitt. On the Need for Caution in Interpreting Test Scores 179

C. P. Hansen and M. J. Reagh. The Predictive Value of Short Intelligence Tests 184

Book Reviews 187

Notes 192

Communications, manuscripts, reviews, summaries, notices, books and magazines for review should be addressed to Dr. James P. Porter, Clark University, Worcester, Mass. Contributors are requested to follow the bibliographical method of citation employed by this Journal, in every case the author, volume and pages should be specified.

THE JOURNAL OF APPLIED PSYCHOLOGY is published quarterly. The subscription price is $4.00 a year. Single numbers $1.25. Remittances and business communications should be addressed to Florence Chandler, Publisher, Clark University, (950 Main St.), Worcester, Mass.

THE
JOURNAL OF APPLIED PSYCHOLOGY

VOL. V JUNE, 1921 No. 2

THE MEASUREMENT OF AGGRESSIVENESS

By H. T. MOORE and A. R. GILLILAND, Dartmouth College

The term aggressiveness, as here used, is synonymous with personal force, initiative, assurance. It is thus understood as standing for that trait which in combination with intelligence and reliability goes far towards completing the essential personal requisites for success. Translated into the terminology of McDougall it is equivalent to an abundance of self-assertion and pugnacity, and a minimum of fear. Described behavioristically it is invariably contrasted with the attitude of inferiority or submission.

While we may never find exactly the same pattern of aggressive behavior in any two individuals, there is nevertheless a widespread and reasonably established belief that one generally knows what to expect of the aggressive man in most of the ordinary circumstances of life. He is more likely to be vigorous, positive, and masterful than the man lacking in this trait; and he is less likely to shrink from notice, to avoid argument, to display a lack of "nerve." Generally speaking his chances of success in any type of undertaking are higher than those of the average man. Undue aggressiveness may handicap personal progress in a certain number of cases, but it is probably not so much the excess of personal force as the lack of tact and sympathy that is the drawback in such cases. So we may say with only slight qualification that, other things being equal, the measure of a man's aggressiveness is the measure of his chance of success.

The laboratory tests of aggressiveness here reported are to some extent the outgrowth of the instinct-emotion tests reported previously in the *American Journal of Psychology*.[1]

[1] "A Method of Testing the Strength of Instincts." *A. J. P.* April 1916, 227-33; and "Laboratory Tests of Fear. Anger and Sex Interest." *A. J. P.* July 1917, 390-5.

But the method of attacking the problem is fundamentally different. The first step in the present experiment was to find two groups of subjects who contrasted to a marked degree in the trait in question. This condition was met by taking from a class of 89 students the highest and lowest fifteen per cent as rated on aggressiveness. This gave thirteen men in each group. Reliable ratings of these men were obtained only after much labor, but it is believed that the method used is quite satisfactory for the present purpose, namely that of guaranteeing a marked difference in aggressiveness between the highest and lowest groups.

The rating of each man was determined by averaging two types of estimates, those made by the faculty and those made by fellow students. In order to secure standard faculty ratings the department of psychology proposed to the faculty a general system of personal ratings of the students. The plan which was adopted is similar in many respects to the officers' ratings used in the army. At the end of the college year each instructor files in the office of the Registrar a rating sheet containing his personal estimates of the men under his instruction. The estimates are made on a scale of five letter grades, as in the case of scholarship marks, and the traits judged are intelligence, aggressiveness, reliability and personality, all of them carefully defined on the rating sheets provided for each instructor. To offset the possibility of perfunctory and therefore careless ratings the point is stressed that no instructor shall under any circumstances submit any ratings which fail to represent a reasonable amount of confidence on his part. It is urged that where it is impossible for a real judgment to be formed, a blank is to be preferred to a forced opinion. Conformably to the above plan the Dartmouth faculty submitted in June 1920 personal data concerning 1,480 students. Practically all the men in the class of 89 students from whom the subjects for the experiment were to be taken were rated by from three to five of their instructors.

Student ratings were similarly obtained from the members of the class. Each student was asked to rate those of his fellow students whom he knew particularly well. From the averages of the faculty and student ratings the list of the thirteen most aggressive and the thirteen least aggressive men was now determined. After making all possible allowance for individual inaccuracies of rating, it seemed absolutely certain that the thirteen highest men, taken as a group, were conspicuously more aggressive for any purpose whatsoever than the lowest thirteen men. As corroborative of this it is

worth mentioning that the thirteen highest men included the president of the senior class, the business manager of the college daily, the manager of track athletics, the quarterback of the football team, the president elect of the Outing Club, and an officer of the Dramatic Association prominent in the production of student plays. The lowest thirteen men were decidedly without prominence in college activities.

To be sure, a certain amount of the difference between the two groups is to be attributed to the factor of intelligence, and indeed this factor can never be completely isolated in any study of personal differences. But an examination of the intelligence ratings and the scholarship marks of the two groups here studied shows that they differ by far less in respect to intelligence than in respect to aggressiveness. If the selection had been of the highest thirteen men in the class on the basis of intelligence ratings, only four of our most aggressive men would have qualified, and one of them would have been included in a list of the thirteen least intelligent. As against this we have the fact that five of the least aggressive were also rated among the thirteen least intelligent, and none of them among the thirteen most intelligent. The average of the scholarship marks of the thirteen most aggressive men is 2.5 on a scale of 4.0 as the maximum. This is about .3 of a point higher than the general average of the college. The average of the thirteen least aggressive men is 1.7, about .5 of a point lower than the college average. There is then a real difference in intelligence in favor of the most aggressive group, almost necessarily so, but it is not to be compared with the difference in aggressiveness. Not one of the highest group had an average rating of less than B on aggressiveness, and not one of the lowest group had an average rating of more than D on the same trait.

Having thus found two groups who differed in nothing so much as in aggressiveness, the next step was to try our tests which seemed calculated to give marked and consistent differences of results for the most and the least aggressive groups. The assumption was that such tests would have considerable significance as affording a measure of the trait. It was hardly to be hoped that any single test would show a point for point correspondence with the existence of the trait,—though the eye movement test does almost approximate such a correspondence. But the expectation was that a grouping of several significant tests would offer a highly suggestive picture of the trait in a given individual,—an aggressiveness profile,—to use Downey's expression.

TABLE I
EYE MOVEMENT TEST

	Name	No. of Eye Movements	
	C.N.	0	
	Z.J.	0	
	A.C.	2	
	R.M.	1	
	N.A.	0	
	J.G.	1	
	E.B.	0	
Most Aggressive	B.W.	0	
	R.W.	1	
	H.C.	0	
	A.J.	0	
	E.L.	1	
	D.R.	0	
			Total 6
	H.M.	7	
	J.J.	1	
	E.R.	4	
	W.P.	4	
	N.F.	4	
	P.G.	5	
Least Aggressive	F.B.	7	
	J.N.	0	
	G.D.	5	
	R.B.	16	
	S.H.	0	
	R.T.	5	
	B.R.	14	
			Total 72

I

EYE CONTROL IN PERSONAL INTERVIEW

Common sense has it that the shifty eye is generally a sure sign of personal weakness, if not of downright dishonesty. The first test was designed with a view to bringing this element of behavior into quantitative relation with the trait aggressiveness. Each of the twenty-six subjects was required to perform a somewhat difficult series of mental additions while constantly returning the fixed gaze of the instructor who sat facing him. The addition series were standardized as to difficulty as described in test II, and the subject performed

them with the knowledge that accuracy and speed were essential. But he was emphatically instructed that under no circumstances should he let his gaze wander from that of the man facing him, as all movements of the eyes were to affect his score seriously.

The details of the addition series are described in the following test. The record for each subject is the total number of eye movements recorded against him during five series of additions. Table I gives the records of the twenty-six subjects.

It will be seen that a total of 72 movements are recorded against the least aggressive as compared with 6 against the most aggressive men. Not one of the aggressive group averted his gaze more than twice during the five series of additions, whereas ten of the thirteen least aggressive subjects shifted their eyes four times or more. If a steady eye does not positively guarantee the presence of aggressiveness, as may be judged by the fact that three of the least aggressive subjects were able to maintain practically constant fixation, a marked lack in this respect is almost invariably accompanied by a lack of aggressiveness.

The question naturally arises whether eye control in personal interview is not perhaps more a matter of "reliability" than of aggressiveness. To test this point the thirteen most reliable and the thirteen least reliable men in the class, as determined by faculty ratings, were submitted to the same experiment. Eight of the thirteen "least reliable" were able to keep within a limit of two eye movements, and the total of the thirteen subjects was 41 movements as compared with 72 for the thirteen least aggressive. Six men were common to the two groups, which would indicate a fairly high correlation between the two traits, but the record for eye movements evidently points to a much more definite relation between eye control and aggressiveness than is the case with reliability.

Whether we compare the average number of movements of the aggressive and the unaggressive subjects, or compare the number of subjects in the two groups who lost control to the extent of making two or more eye movements, we find evidence of at least ten times as much control in the upper group as in the lower. Thus the simple behavioristic fact of the ability to look another person in the eye seems to have such a high significance regarding the presence or absence of aggressiveness as to warrant giving it an extremely prominent place in any scoring method devised as a measure of this trait. The correspondence is in fact so close as to justify the generalization that a stop watch and a pair of fixed eyes are the

only indispensable laboratory equipment necessary for estimating roughly the degree of aggressiveness in at least four-fifths of the subjects.

II

Fear Distraction Tests

The purpose of the distraction tests was to determine how far the ability to resist a type of distraction involving a certain element of fear is related to the possession or lack of aggressiveness. Preliminary to all of these tests each subject was required to practice fifty times performing mental additions according to a uniform plan involving constant increments of one, through a series of nine additions. If such a series began with 36, it would continue; 37, 39, 42, 46, 51, 57, 64, 72, 81. If it began at 29, it would continue; 30, 32, 35, 39, 44, 50, 57, 65, 74. Thus a series might begin with any number and would consist of nine additions, the numbers added being; 1, 2, 3, 4, 5, 6, 7, 8, and 9. On the day prior to the experiment each subject had practiced such additions with fifty series. This amount of practice seemed to bring each man reasonably near to the limit of his natural ability at such a task. The thirteen most aggressive subjects were able after this amount of practice to perform ten series of additions under normal conditions at an average time of 12.6 seconds. The thirteen least aggressive subjects averaged 15.6 seconds for the same ten series under normal conditions. The median of the thirteen mean variations for the normal series of the aggressive subjects was 2.0 seconds; that for the least aggressive subjects was 2.3 seconds. The difference in the normal variability of the two groups was thus negligibly small. In the comparison of any two individuals the factor of normal variability in the addition series will have to be taken into account, but for purposes of comparing these two groups as a whole it is of less importance.

Distraction by Staring

Immediately after the ten normal series were completed the subjects were asked to perform the same task under conditions of slight emotional distraction. The first of these distractions was that of having to meet the fixed gaze of the instructor facing the subject, as described in the eye movement test. This distraction record is in fact the time record which was obtained simultaneously with the eye movement score. While one experimenter counted the eye movements which gave the

record for Table I, another timed the five series of additions performed under these conditions. The average increase of time as compared with the normal average is obtained as the measure of the subject's distraction. Table II gives the results for this test.

TABLE II

DISTRACTION TEST—STARING

	Name	Time Increase With Test	A.D. Without Distraction	No. of Cases in which Distraction Time is Greater Than A.D.	
Most Aggressive	C.N.	—1.9	.6	...	
	Z.J.	.4	2.3	...	
	A.C.	.4	2.8	...	
	R.M.	— .1	2.0	...	
	N.A.	3.0	1.3	1	
	J.G.	— .4	2.5	...	
	E.B.	5.1	1.2	1	
	B.W.	— .1	3.5	...	
	R.W.	0.	.8	...	
	H.C.	.5	2.8	...	
	A.J.	.9	2.5	...	
	E.L.	2.9	1.0	1	
	D.R.	—1.4	1.3	...	
					Total 3
Least Aggressive	H.M.	9.7	2.5	1	
	J.J.	1.4	3.5	...	
	E.R.	5.3	2.7	1	
	W.P.	2.5	2.1	1	
	N.F.	— .5	2.3	...	
	P.G.	— .6	1.8	...	
	F.B.	4.9	2.4	1	
	J.N.	—2.0	2.6	...	
	C.D.	— .3	1.5	...	
	R.B.	2.8	1.4	1	
	S.H.	4.8	1.5	1	
	R.T.	5.0	4.2	1	
	B.R.	8.9	2.0	1	
					Total 8

The most aggressive subjects are listed in the upper half of the table, and the least aggressive subjects in the lower half. The first column of numbers shows the average amount of time increase of each subject under the distraction conditions, as measured by comparing the average time for the five staring tests with the average time for the ten normal trials. When

the time made in spite of distraction was more rapid than normal, this is indicated by a minus sign before the number. The middle column gives the average deviations for each subject in the normal series, and the right hand column indicates the cases in which the amount of time increase under distracting conditions was greater than the average deviation under normal conditions. The results may be summarized as follows: for the aggressive subjects the staring caused an average delay of .4 of a second; for the unaggressive subjects it caused an average delay of 3.2 seconds. Only three of the aggressive subjects were delayed by an amount equal to their average deviations under normal conditions, whereas eight of the unaggressive subjects have a distraction time greater than the A. D. The results indicate therefore approximately three times as much probability that an unaggressive person when stared at will be deterred to a considerable degree than that an aggressive person will be thus delayed.

Distraction by Electric Shock

In the second of the three distraction tests the element of distraction was the expectation of an electric shock which was to come during or at the end of each of the five series of additions. The subject was told that he would receive shocks, which might be of any intensity from 75 to 220 volts, and which might be expected at any time during or between the addition series. Actually he was not given more than 75 volts at any time, and the shock was always given just as he was finishing his last number in a series, so that the distraction was almost wholly imaginary rather than sensory.

The accompanying picture shows the apparatus used in the shock test.

The base of the apparatus consists of a 9 by 18 inch board, divided into two square sections by a board upright 5 inches high. An electric light plug and five sockets, fitted with 60 watt lamps, are on one side of the partition. The sockets are interconnected by a combination of series and parallel wiring. On the other side of the partition are a small voltmeter and two binding posts. The voltmeter and the posts are connected to the lamp sockets by four wires running through the partition. Two brass hand electrodes are connected one to each binding post. By tightening or loosening certain lamp bulbs a difference of potential of 75 or 220 volts can be sent either through the meter, or to the electrodes, or to both. The lamps, of course, are lighted whenever they are thrown in the circuit,

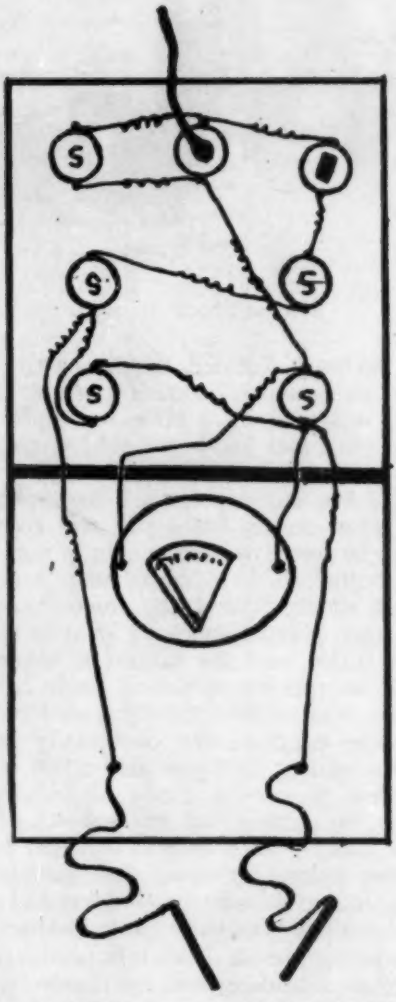
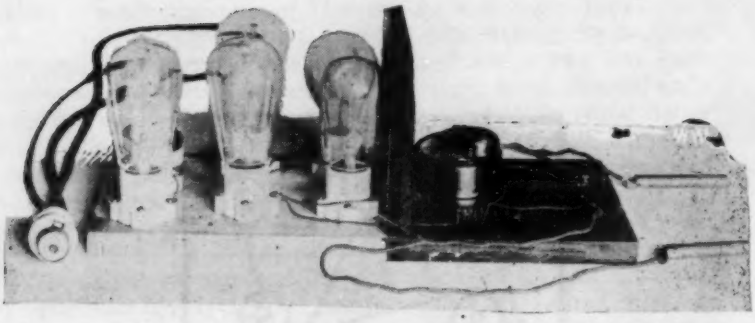


Diagram Showing Wiring for the Electric Shock Apparatus.



Electric Shock Apparatus

but they are partially shielded by the partition, and carbon lamps are used so as not to produce a bright distracting light.

The subject is seated by a table on which the apparatus rests. The experimenter hands the subject the electrodes and says: "This is an electric apparatus. As you begin your additions I shall begin manipulating these light bulbs. When I loosen or tighten certain bulbs you will receive an electric shock. It may be weak or strong. It is not dangerous, but may be very uncomfortable. Just to show how it feels I will give you a weak shock. Watch the pointer on the voltmeter." The subject is then given a shock of 75 volts. The electrodes are laid on the table, and the subject is asked to watch the voltmeter while another manipulation sends 220 volts through it. He is again told to take the electrodes in hand, and the test begins. The experimenter continually manipulates the bulbs during the adding, and just at the last addition in each series the subject receives a shock of 75 volts. No shock is given during the adding, but the subject is continually in a state of expectancy. Each subject is bound by a promise of secrecy regarding everything about the experiment. The tests were completed in two days and no subject had any knowledge further than that there was to be some kind of electric shock.

The results are given in Table III, and according to the tabular scheme already described for Table II.

The figures for time increases in the left hand column show that five of the aggressive subjects resisted the distraction with complete success, whereas all of the twelve, unaggressive subjects were positively affected to the extent of an increase of 2.5 seconds or more. The average shock-delay of the least aggressive subjects was 6.1 seconds as compared with

TABLE III

DISTRACTION TEST—SHOCK

	Name	Time Increase With Test	A.D. Without Distraction	No. of Cases in which Distraction Time is Greater Than A.D.	
Most Aggressive	C.N.	2.9	.6	1	
	Z.J.	2.5	2.3	1	
	A.C.	10.0	2.8	1	
	R.M.	3.0	2.0	1	
	N.A.	4.8	1.3	1	
	J.G.	5.8	2.5	1	
	E.B.	3.9	1.2	1	
	B.W.	—2.3	3.5	...	
	R.W.	.4	.8	...	
	H.C.	0	2.8	...	
	A.J.	— .9	2.5	...	
	E.L.	8.0	1.0	1	
	D.R.	— .4	1.3	...	Total 8
Least Aggressive	H.M.	10.1	2.5	1	
	J.J.	9.9	3.5	1	
	E.R.	6.0	2.7	1	
	W.P.	17.1	2.1	1	
	N.F.	3.3	2.3	1	
	P.G.	...	1.8	...	
	F.B.	2.5	2.4	1	
	J.N.	2.6	2.6	1	
	G.D.	4.7	1.8	1	
	R.B.	5.5	1.4	1	
	S.H.	3.8	1.5	1	
	R.T.	2.9	4.2	...	
	B.R.	4.7	2.0	1	Total 11 (One score missing)

an average of 2.2 seconds for the aggressive group. The individual variations among the aggressive subjects seem to indicate that this test has a less specific relation to aggressiveness than does the staring test, and this is somewhat borne out by the fact that when the same test was given to the thirteen most reliable and the thirteen least reliable men in the class, the difference was by far more striking than for the aggressiveness groups. Only one of the thirteen "most reliable" men had a shock-delay of more than three seconds, whereas eleven of the thirteen "least reliable" men had delays upwards of three seconds. However, the fact that the ability to with-

stand shock distraction is possessed by 40 per cent of the aggressive subjects and by none of the unaggressive ones justifies the inclusion of this test as having considerable symptomatic value regarding the trait.

Distraction by Snake

In the third test the method of distraction was to place a dead snake, suitably coiled, and pinned to a cork board, about ten inches in front of the face of the subject during one-half of his adding. Only one series of additions was recorded for this test, as it was found that adaptation was very rapid. This means, of course, that the results have a larger probable error than those recorded for the other types of distraction. The detailed record is given in Table IV.

TABLE IV

Distraction Test—Snake

	Name	Time Increase With Test	A.D. Without Distraction	No. of Cases in which Distraction Time is Greater Than A.D.	
Most Aggressive	C.N.	9.1	.6	1	
	Z.J.	4.4	2.3	1	
	A.C.	6.6	2.8	1	
	R.M.	2.7	2.0	1	
	N.A.	3.6	1.3	1	
	J.G.	10.6	2.5	1	
	E.B.	7.7	1.2	1	
	B.W.	4.3	3.5	1	
	R.W.	1.4	.8	1	
	H.C.	— .5	2.8	...	
	A.J.	4.3	2.5	1	
	E.L.	5.9	1.0	1	
	D.R.	0.	1.3	...	Total 11
Least Aggressive	H.M.	9.7	2.5	1	
	J.J.	11.4	3.3	1	
	E.R.	12.6	2.7	1	
	W.P.	.1	2.1	...	
	N.F.	1.1	2.3	...	
	P.G.	9.8	1.8	1	
	F.B.	21.7	2.4	1	
	J.N.	.8	2.6	...	
	G.D.	8.7	1.8	1	
	R.B.	11.6	1.4	1	
	S.H.	3.2	1.5	1	
	R.T.	7.0	4.2	1	
	B.R.	9.5	2.0	1	Total 10

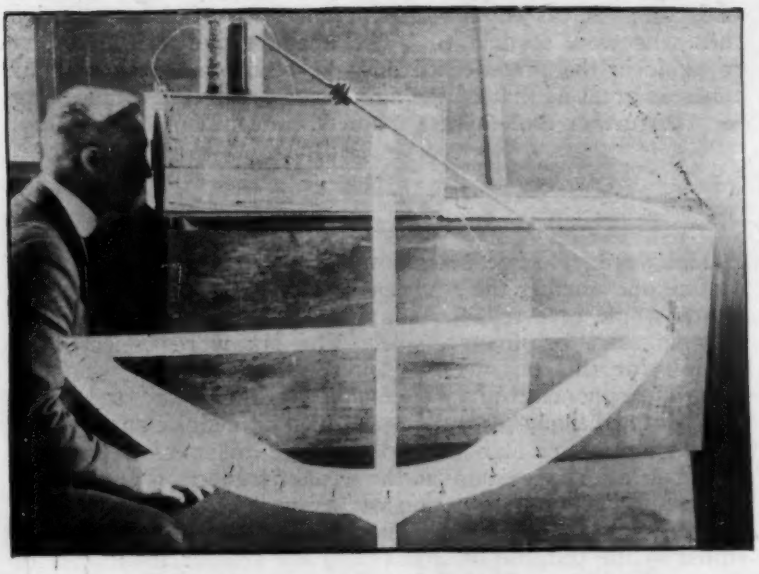
The time increase for practically all of the subjects was greater than the average deviation in the normal series, but the average delay of 8.2 seconds for the unaggressive subjects was practically double the 4.6 seconds for the aggressive ones. The most noticeable fact here was the greater liability of the unaggressive subjects to become somewhat panicky. Generally a delay of nine seconds or more was accompanied by very definite movements of withdrawal or avoidance, and it will be seen from the figures in the first column of the table that eight of the unaggressive men were delayed by nine seconds or more, as compared with only two of the aggressive men who were so delayed. The unaggressive man's liability to panic in the presence of a snake seems to be about four times as great as that of the aggressive man. It is interesting to note further that this test showed no relation to reliability.

III

Word Association Tests

The word association tests were designed with a view to finding out whether the study of responses to selected stimulus words would give reliable data for the differentiation of men as aggressive or unaggressive. The six words selected as stimuli for this test were: 'enterprise,' 'success,' 'danger,' 'death,' 'opponent,' and 'company.' 'Enterprise' and 'success' seemed likely to contain positive and definite suggestions for the aggressive man, and 'death' and 'danger' to be more full of definite suggestions to the weaker subjects. 'Opponent' offered the concrete situation of personal contest, and 'company,' being adapted to so many possible meanings, seemed suited to the purpose of uncovering the readier tendencies in the mind of the subject.

Apparatus. The use of a lip key or similar recording device in circuit with a chronoscope was soon found unsatisfactory on account of the frequency with which the subjects tended to make lip movements, or even to begin to vocalize, before they were actually ready with a response. Whenever this occurred, it meant not only that the time record was spoiled, but that the stimulus word could never be given to that subject again under the original conditions. Moreover, oral presentation of words proved somewhat unsatisfactory on account of the possibility of slight misunderstandings, and the slight variations in emotional suggestion on the part of the experimenter. The apparatus finally constructed is shown in the accompanying picture.



Word Exposure and Association Time Apparatus

It consists of two main parts, the timing apparatus and the exposure box. The first of these consists of a standard, a pendulum, and a graduated metal arc. The standard is 42 inches high and supported by a base. The pendulum is a metal rod, 40 inches long, and is attached 10 inches from one end to the top of the standard. The shorter end of the pendulum is weighted so that it will make a half swing in exactly one second. The metal arc is one-fifth the length of the circumference of the pendulum swing. It is attached to the standard, one half on either side, just back of the path of the end of the pendulum. The arc is graduated so that time may be measured to within 5 hundredths of a second from the pendulum swing.

The exposure box is 10x12x20 inches and is closed except for a 2x4 inch opening at one end and a 1x2 inch opening at the other. From a revolving spool, attached just above the smaller opening, a sheet of paper 2 inches wide and 70 inches long extends to another similar spool just under the opening. The paper may be wound from one spool to another by means of handles attached to the spools. A piece of metal fits over the paper in front of the opening of the box so that it excludes all light. The stimulus words are arranged about an inch apart on the strip of paper, and by proper adjustment of the spools any word may be made ready for the exposure which is to result from the closing of the light circuit. For this purpose two 6 candle power lights are provided, one at either side of the larger opening of the box. These are wired in series with a battery and connected with the catch which suspends the pendulum in such a way that the moving of the catch to release the pendulum closes the light circuit, and thus exposes the stimulus word to view.

The subject sits with his eyes close to the larger opening of the exposure box, looking towards the paper about to be exposed at the other end of the box when the light connection is made. When the lights are not on, the box is dark, and the word on the paper is therefore not visible. The subject is given instructions to look for the word that is to be exposed, and to respond as quickly as he can with the first word that he can associate with it. A second experimenter watches the pendulum, and by noting the position of the pendulum on the graduated arc at the moment of the response, he is able to record the time to within 5 hundredths of a second.

The two sources of error in determining word reaction time by this method are: (1) the auditory reaction time of the second experimenter and (2) the errors in reading the swing

of the pendulum from the graduations of the arc. The first of these is probably very nearly constant for the same experimenter throughout a series of observations. The latter error may be in either direction but is seldom more than 5 hundredths of a second. In general the time records here presented are a small fraction of a second too long. But as this tendency is constant in all the records, and the method offers the subject no possible way of stealing a march on the experimenter, it seems highly satisfactory.

The results of the association experiment are set forth in Table V.

The six stimulus words are indicated in the horizontal line at the top of the table. Each subject's responses to the six words are given after his initials. Just above each of the response words is a large number on the right which indicates the time record in hundredths of a second. The small number just above the response word shows the frequency with which that response was given when the experiment was repeated with 78 subjects. Twenty-two of the original twenty-six subjects were available for this experiment. The upper part of the table contains the results for the eleven very aggressive subjects, and the lower half the results for the eleven least aggressive ones.

There is evidently a significant difference between the two groups in the character of the words chosen. In response to the words 'enterprise' and 'success' the aggressive subjects are almost always definite and positive. Of their twenty-one responses fourteen are decidedly colorful;—'initiative,' 'push,' 'money' twice, 'activity,' 'scheme,' 'undertake,' 'ambition,' 'power,' 'gain,' 'win' twice, 'wealth' and 'advance.' Five responses were sufficiently general to be classified as of doubtful significance;—'building,' and 'business' four times. One was apparently based on verbal similarity;—'success'—'successful.' Only one response;—'failure' was negative. The twenty-two responses of the eleven least aggressive men to the same two words were strikingly different. Only three were positive and colorful;—'work' and 'ambition,' the latter occurring twice. Twelve are doubtful, colorless, or passive;—'physics,' 'American Magazine,' 'oil,' 'friends,' 'good,' 'college,' 'happiness,' 'business' four times, and 'industry.' The negative response;—'failure' occurred four times. Briefly, there is four times as much probability of a definite, forward-looking response to 'enterprise' and 'success' from the aggressive man as from the unaggressive. Furthermore, there is

TABLE V
MOST AGGRESSIVE SUBJECTS

	Stimulus Words					
	Company	Enterprise	Opponent	Success	Danger	Death
Name C.N.	5 400 army	1 140 initiative	1 260 box	1 290 power	34 150 fear	1 275 cotton
Z.J.	1 220 battery	2 175 building	26 135 enemy			1 135 living
A.C.	3 200 girl	22 105 business	1 255 politics	2 170 money	2 95 dread	33 175 life
R.M.	1 190 friendly	1 190 push	2 400 hatred	2 135 gain	34 100 fear	33 130 life
N.A.	3 200 girl	2 140 money	2 175 opposite	3 200 win	34 135 fear	3 140 die
J.G.	2 130 people	1 130 activity		1 135 wealth	34 135 fear	11 230 fear
E.B.	2 200 alone	1 310 scheme	26 200 enemy	3 210 win	2 150 dread	33 115 life
R.W.	1 130 association	2 100 undertake	7 95 foe	1 200 successful	34 95 fear	2 135 death
H.C.	2 210 business	22 95 business	1 300 war	1 250 advance	34 175 fear	11 340 fear
A.J.	1 130 insurance	22 130 business	26 195 enemy	5 140 business	1 200 car	33 155 life
E.L.	2 135 society	4 180 ambition	1 130 adversary	25 135 failure	8 150 safety	133 100 life
	LEAST AGGRESSIVE SUBJECTS					
H.M.	1 145 good	22 145 business	26 155 enemy	1 420 Amer. Mag.	2 210 red flag	2 155 grave
J.J.	2 150 people	1 160 enterprise	26 160 enemy	2 130 happiness	8 200 safety	33 140 life
E.R.	2 120 group	4 140 ambition	9 160 fight	2 200 succeed	2 400 death	11 140 fear
W.P.	3 100 girl	1 280 oil	2 85 football	1 140 college	1 135 money	1 115 killed
N.F.	2 115 alone	1 495 encounter	2 100 antagonist	5 300 business	2 430 act	2 130 birth
F.B.	2 200 man	22 235 business	3 200 man	2 240 good	2 150 red flag	1 150 man
J.N.	5 200 army	1 400 work	26 200 enemy	25 95 failure	1 320 enemy	33 160 life
G.D.	9 145 friend	3 135 industry	9 150 friend	25 150 failure	1 300 fly	33 100 life
R.B.	6 230 friends	22 230 business	26 140 enemy	1 180 friends	34 95 fear	3 285 funeral
S.H.	6 105 friends	1 250 failure	8 130 friend	25 115 failure	34 145 fear	33 120 life
R.T.	1 400 Ford	4 330 ambition	1 300 hardware	1 240 physics	34 160 fear	1 190 accident

only one fourth as much probability of his giving a negative response, such as 'failure.'

Almost as clear cut is the difference in the character of the responses of the two groups to the words 'danger' and 'death.' In this case the aggressive responses are mostly very general and commonplace, while those of the unaggressive are more concrete and vivid. Of the twenty-one responses of the aggressive subjects only three seem to indicate anything resembling a vivid, personal experience. These words were 'car' and 'dread,' the latter occurring twice. The response 'cotton' to the stimulus word 'death' is doubtful, and the remaining seventeen are either so common as to be meaningless, or else are mere verbal variations of the stimulus word;—'life' five times, 'fear' seven times, 'safety,' 'die,' 'death.' By contrast, the unaggressive subjects present at least 10 out of 22 responses which are colorful and suggestive of personal interpretation;—'money,' 'fly,' 'act,' 'enemy,' 'red flag' twice, 'grave,' 'killed,' 'funeral,' 'accident.' The remaining twelve are either doubtful or merely verbal;—'life' four times, 'fear' four times, 'safety,' 'birth,' 'man,' and 'death.' In short, the probability of a definite and vivid response to the words 'danger' and 'death' is four times as great for the unaggressive man as for the aggressive man.

The responses to 'company' show exactly twice as many military, sexual, and commercial associations for the aggressive as for the unaggressive subjects. 'Army' and 'battery' in the aggressive list are to be compared with 'army' once in the lower list; 'girl' occurs twice in the upper list to once in the lower list; and 'business' and 'insurance' in the upper list are to be compared with 'Ford' in the lower list. The evidence, which is slight, points to the conclusion that aggressiveness doubles the probability of a definite and energetic response to the word 'company.'

The responses to the word 'opponent' are extremely difficult to interpret, and seem to show such an influence of the language habit that the word is probably not a suitable one for giving scorable results. The aggressive responses;—'box,' 'politics,' 'hatred,' and possibly 'war' are more pointed expressions of personal antagonism than any of the words in the unaggressive list except 'fight;' but 'football' and 'hardware' leave considerable doubt, and both lists contain a high per cent of merely verbal associations.

Confining our statement of the results to the five words which afforded the most definite responses, we may say that the very aggressive man is four times as likely to be positive

and definite in his responses to 'enterprise' and 'success,' twice as likely to give an energetic type of response to 'company;' and only one-fourth as likely to respond definitely and vividly to 'danger' and 'death' as the very unaggressive man. These differences are large enough to warrant the use of the word responses as one of the minor scoring elements in a measurement of aggressiveness.

The time records for the responses show a marked difference for the stimulus words 'danger' and 'enterprise.' Otherwise the average differences for the aggressive and the unaggressive groups are less than half a second. For 'danger' the unaggressive subjects take an average time of 2.31 seconds as compared with an average time of 1.26 seconds for the aggressive subjects. For 'enterprise' the unaggressive subjects average 2.55 seconds and the aggressive subjects 1.54 seconds. Failures to respond are not here reckoned in averaging the time, for the reason that such failures were due to imperfect exposure of a few words in the early part of the experiment.

Seven of the unaggressive responses to 'enterprise' required two seconds or more, as compared with one such case among the aggressive subjects. Likewise six of the unaggressive subjects required two seconds or more to respond to the word 'danger,' as compared with one such case among the aggressive subjects. This difference is not due to the greater average quickness of response of the aggressive subjects, as their averages for the remaining four words are actually about a fifth of a second slower than the corresponding average for the unaggressive subjects. Apparently it is due to some peculiar relation of the two words to the two classes of subjects. The obvious assumption in regard to 'danger' is that it comes near to awakening true emotional disturbances, much more nearly so in the unaggressive subjects. Their slowness in responding to 'enterprise' might be due either to emotional inhibition or to paucity of ideational material, either a matter of self reproach or of an unfamiliar concept. Whatever the theoretical explanation, the time record of the two groups of subjects is large enough and constant enough to justify its use in differentiating the two types of individuals.

Conclusion

We are now ready to propose a method of scoring aggressiveness on the basis of our tests. In order to have a perfect record on the trait, a man should be able to maintain perfect control of his eyes in the interview test; he should withstand

with complete success all the three types of distraction; and in the association tests he should respond definitely and positively to 'enterprise,' 'success,' and 'company,' and indefinitely to 'danger,' and 'death;' moreover, his reaction time to 'enterprise' and 'danger' should not be noticeably longer than his normal word reaction time. The scoring method proposed makes deductions from 100 points for each failure of a subject to meet satisfactorily the four types of requirement just mentioned. Corresponding to the relative importance of the different tests for this purpose a maximum possible deduction of 50 points is allowed for eye movements in test I, a maximum of 25 points each for the distraction and association tests. Stated more in detail, the scheme of deductions for each type of failure is as follows: In the interview test 5 points are deducted for every movement up to ten movements. The maximum deduction is therefore 50 points. For the staring distraction record two points are deducted for each second of time increase in excess of three seconds above his normal adding time. The maximum deduction allowed is 10 points. For the shock distraction 2 points are deducted for every second of time increase in excess of five seconds. The maximum deduction allowed is 10 points. For the snake distraction test 1 point is deducted for each second of time increase in excess of seven seconds. The maximum deduction allowed is 5 points. In the word association test 2 points are deducted for each negative response to 'enterprise,' 'success,' or 'company' and for each vivid, personal type of response to 'death' or 'danger.' One point is deducted for each colorless or doubtful response to 'enterprise,' 'success' or 'company.' The maximum deduction on the basis of the word content of responses is thus 10 points. For the time of response to 'enterprise' and 'danger' 1 point is deducted for every fifth of a second in excess of two seconds. The maximum deduction allowed on this account is 15 points.

The score of any individual subject in aggressiveness is to be obtained by totaling his deductions and subtracting the amount from 100 as shown in table VI. On this basis the scores of our thirteen most aggressive subjects are as follows: C. N. 97, Z. J. 100, A. C. 79, R. M. 94, N. A. 98, J. G. 86, E. B. 83, B. W. 100 (lacking word test), R. W. 93, H. C. 98, A. J. 97, E. L. 84, D. R. 100 (lacking the word test). The scores for the thirteen least aggressive subjects are as follows: H. M. 35, J. J. 79, E. R. 55, W. P. 60, N. F. 59, P. G. 72 (lacking word test), F. B. 36, J. M. 81, G. D. 63, R. B. 37, S. H. 88, R. T. 61, B. R. 38 (lacking word test).

TABLE VI
FINAL SCORES OF 26 SUBJECTS
Most Aggressive

	Eye Movement	Staring	Shock	Snake	Words	Time	Score
C.N.	..	..	..	2	1	..	97
Z.J.	..	..	..	.	.	..	100
A.C.	10	..	10	.	1	..	79
R.M.	5	..	..	.	1	..	94
N.A.	..	2	..	.	.	..	98
J.G.	5	..	2	4	3	..	86
E.B.	..	6	..	1	4	6	83
B.W.	..	..	..	.	x	x	100
R.W.	5	..	..	.	2	..	93
H.C.	..	..	..	.	2	..	98
A.J.	..	..	..	.	3	..	97
E.L.	5	2	6	.	3	..	84
D.R.	..	..	..	.	x	x	100
			Least	Aggressive			
H.M.	35	10	10	3	6	1	35
J.J.	5	..	10	4	2	..	79
E.R.	20	4	2	5	4	10	55
W.P.	20	..	10	.	6	4	60
N.F.	20	..	..	.	6	15	59
P.G.	25	..	..	3	x	x	72
F.B.	35	4	..	5	5	15	36
J.N.	..	..	..	.	4	15	81
G.D.	25	..	..	2	5	5	63
R.B.	50	..	2	5	4	2	37
S.H.	..	4	..	.	5	3	88
R.T.	25	4	..	.	3	7	61
B.R.	50	10	..	2	x	x	38

The average score of the aggressive group is 93 and that of the unaggressive, is 58.8. None of the aggressive subjects score less than 79, whereas nine of the least aggressive make less than 64, and one of the others is incomplete. Three of the twenty-six subjects, J. N., S. H., and J. J. appear as noticeable exceptions. It is the opinion of the writers, based on considerable acquaintance with the subjects, that the error in regard to at least two of them lies in original personal ratings rather than in the finding of the test. In any event we seem to be justified in stating that there is hardly one chance in twenty-five that a man weak in aggressiveness would score as high as 85; and there is almost no chance that a very aggressive person will score lower than 70. The writers believe that this test approximates a true measurement of aggressiveness more nearly than does the Army Alpha examination approximate the measurement of intelligence.

Moreover, the test can be given in modified and abbreviated form with no further equipment than a stop watch and a set of addition tables, with the additions all made according to the specifications in test II. In this case it would be necessary to omit the shock and snake distraction tests; also to give the words for the association test orally, and to time the responses with a stop watch. The loss of two of the 10 point tests can be compensated for in the scoring by adding 25% to the total of the deductions made on the basis of the remaining tests.

Aside from the time required of each individual subject for practicing his 50 additions with the tables, the actual time of administering the abbreviated form if the test should not be more than three minute. The longer form with apparatus requires about seven minutes per subject.

THE BEDAUX-UNIT PRINCIPLE OF INDUSTRIAL MEASUREMENT

Its Origin, Its Laws and Applications

By Charles E. Bedaux

(This article is copyrighted by the author, and reproduction in any form is forbidden except by his express permission.)

Ever since industry reached the tremendous proportions afforded by the mechanical era, the need has been felt for a system of measurements permitting the exact determination of the quantitative value of the human physical effort.

It is due to the lack of such a system that to-day we meet deeply set differences between management and labor. Either side ignoring the value of a given effort made it possible for the employee to practice, undetected, a restraint in production, while the employer, in some instances, protected by the same ignorance, has been able to arbitrarily decrease the money value of the operation performed.

A correct system of measurement of the industrial effort has many applications:

1. It permits a protective and adequate system of remuneration;
2. It gauges the capacity of the human being or of a group of employees comprising one department, permits an exact comparison between departments, between plants, and automatically indicates temporary or permanent conditions that are conducive to failure;
3. It may be used as a basis for apportionment of the overhead burden;
4. It allows an automatic and constant comparison between the standard and actual labor cost;
5. It renders scientific planning and scheduling of manufacture possible;
6. It is the only foundation on which any form of labor representation in shop management may stand.

The exponents of the piece-work system have recognized that the basic weakness of their method lies in the inability to guarantee the rate set, as it is based on the monetary value of labor, a commodity which fluctuates. To guarantee the

rate is to expose the manufacturer to failure, in case of a general decrease in the labor market. To cut the rate destroys the confidence of the working man and compels a restraint in production.

The monetary value of labor, being a variable, cannot be used as a means to measure industrial effort.

Because of fluctuations in this value, not only between periods of depression or growth, but also within each period, among the various classes of labor, it has been impossible for the manufacturer to gauge the respective productive capacities of his various departments, or of his entire plant, in comparison with a competitive or co-operative institution.

The result has been that many plants, sincerely believing that their productivity rate was high, have failed because of their error, the detection of which was impossible without an adequate system of measurement of the human effort applied to industry.

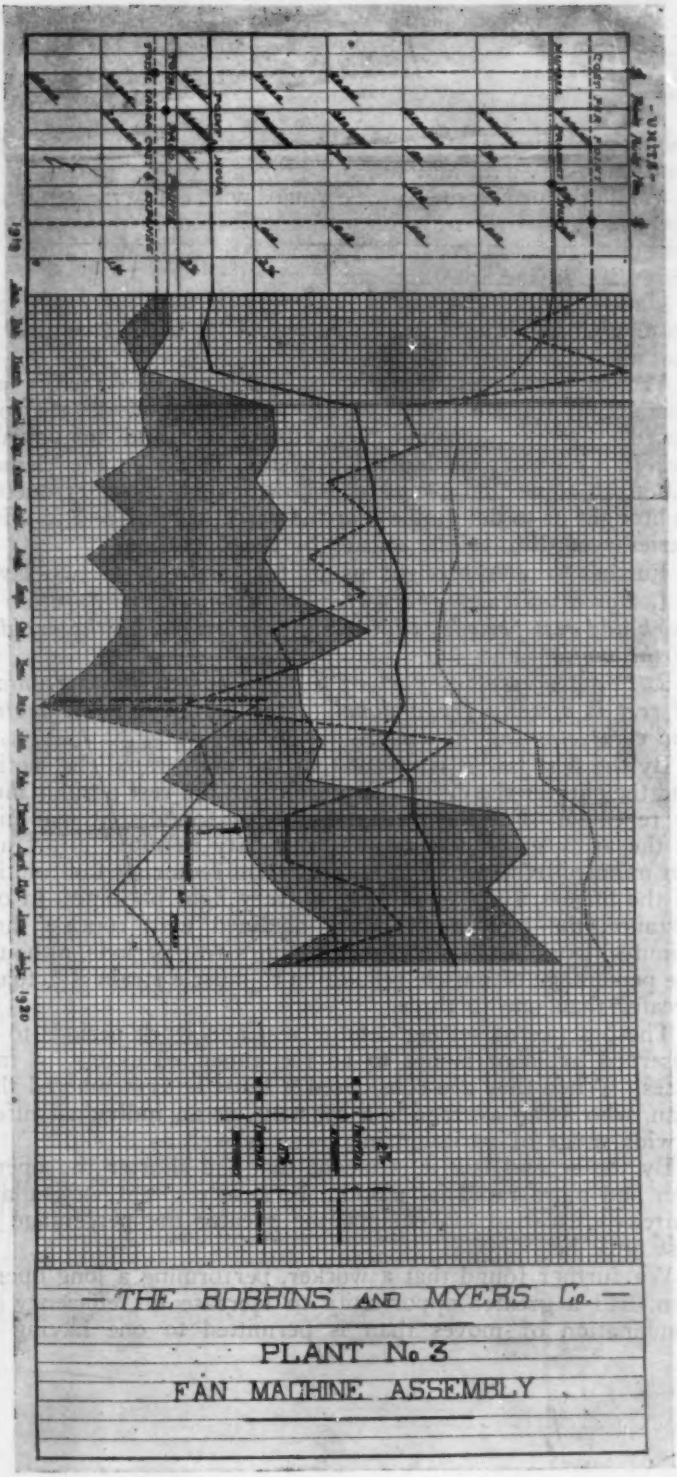
It is also because of the absence of such a system that the monetary value of productive labor has been recorded for cost purposes by the taking of the time of the beginning and the ending of each operation. This performance is generally inaccurate, or, when accurate, necessitates waiting time for recording—a waste of labor, both in the productive and in the clerical forces.

To that source also must be traced the erroneous distribution, too often met, of the overhead burden, wherein an operation costing one dollar ($1.00) in productive labor, performed on a machine worth one thousand dollars ($1,000.00) and depreciating in twenty years, for instance, will be charged with the same amount of overhead as an operation also costing one dollar ($1.00) in productive labor, but performed on a machine the cost of which is one hundred thousands dollars ($100,000.00) with a rate of depreciation that must be absorbed in ten years.

Because of the lack of an adequate system of measurement, incapable operators have been kept in positions that have been refused to men who were best indicated to hold them; promotion has been arbitrary and illogical, thus entailing losses to both capital and labor and creating a state of mind on either side which is wholly destructive.

Part Two

For many years we have seen the need of a unit of measure and our study first led us to determine that such a system could not be based on monetary values which fluctuate. We



— THE ROBBINS AND MYERS Co. —

PLANT No. 3

FAN MACHINE ASSEMBLY

only one fourth as much probability of his giving a negative response, such as 'failure.'

Almost as clear cut is the difference in the character of the responses of the two groups to the words 'danger' and 'death.' In this case the aggressive responses are mostly very general and commonplace, while those of the unaggressive are more concrete and vivid. Of the twenty-one responses of the aggressive subjects only three seem to indicate anything resembling a vivid, personal experience. These words were 'car' and 'dread,' the latter occurring twice. The response 'cotton' to the stimulus word 'death' is doubtful, and the remaining seventeen are either so common as to be meaningless, or else are mere verbal variations of the stimulus word;—'life' five times, 'fear' seven times, 'safety,' 'die,' 'death.' By contrast, the unaggressive subjects present at least 10 out of 22 responses which are colorful and suggestive of personal interpretation;—'money,' 'fly,' 'act,' 'enemy,' 'red flag' twice, 'grave,' 'killed,' 'funeral,' 'accident.' The remaining twelve are either doubtful or merely verbal;—'life' four times, 'fear' four times, 'safety,' 'birth,' 'man,' and 'death.' In short, the probability of a definite and vivid response to the words 'danger' and 'death' is four times as great for the unaggressive man as for the aggressive man.

The responses to 'company' show exactly twice as many military, sexual, and commercial associations for the aggressive as for the unaggressive subjects. 'Army' and 'battery' in the aggressive list are to be compared with 'army' once in the lower list; 'girl' occurs twice in the upper list to once in the lower list; and 'business' and 'insurance' in the upper list are to be compared with 'Ford' in the lower list. The evidence, which is slight, points to the conclusion that aggressiveness doubles the probability of a definite and energetic response to the word 'company.'

The responses to the word 'opponent' are extremely difficult to interpret, and seem to show such an influence of the language habit that the word is probably not a suitable one for giving scorable results. The aggressive responses;—'box,' 'politics,' 'hatred,' and possibly 'war' are more pointed expressions of personal antagonism than any of the words in the unaggressive list except 'fight;' but 'football' and 'hardware' leave considerable doubt, and both lists contain a high per cent of merely verbal associations.

Confining our statement of the results to the five words which afforded the most definite responses, we may say that the very aggressive man is four times as likely to be positive

and definite in his responses to 'enterprise' and 'success,' twice as likely to give an energetic type of response to 'company;' and only one-fourth as likely to respond definitely and vividly to 'danger' and 'death' as the very unaggressive man. These differences are large enough to warrant the use of the word responses as one of the minor scoring elements in a measurement of aggressiveness.

The time records for the responses show a marked difference for the stimulus words 'danger' and 'enterprise.' Otherwise the average differences for the aggressive and the unaggressive groups are less than half a second. For 'danger' the unaggressive subjects take an average time of 2.31 seconds as compared with an average time of 1.26 seconds for the aggressive subjects. For 'enterprise' the unaggressive subjects average 2.55 seconds and the aggressive subjects 1.54 seconds. Failures to respond are not here reckoned in averaging the time, for the reason that such failures were due to imperfect exposure of a few words in the early part of the experiment.

Seven of the unaggressive responses to 'enterprise' required two seconds or more, as compared with one such case among the aggressive subjects. Likewise six of the unaggressive subjects required two seconds or more to respond to the word 'danger,' as compared with one such case among the aggressive subjects. This difference is not due to the greater average quickness of response of the aggressive subjects, as their averages for the remaining four words are actually about a fifth of a second slower than the corresponding average for the unaggressive subjects. Apparently it is due to some peculiar relation of the two words to the two classes of subjects. The obvious assumption in regard to 'danger' is that it comes near to awakening true emotional disturbances, much more nearly so in the unaggressive subjects. Their slowness in responding to 'enterprise' might be due either to emotional inhibition or to paucity of ideational material, either a matter of self reproach or of an unfamiliar concept. Whatever the theoretical explanation, the time record of the two groups of subjects is large enough and constant enough to justify its use in differentiating the two types of individuals.

Conclusion

We are now ready to propose a method of scoring aggressiveness on the basis of our tests. In order to have a perfect record on the trait, a man should be able to maintain perfect control of his eyes in the interview test; he should withstand

with complete success all the three types of distraction; and in the association tests he should respond definitely and positively to 'enterprise,' 'success,' and 'company,' and indefinitely to 'danger,' and 'death;' moreover, his reaction time to 'enterprise' and 'danger' should not be noticeably longer than his normal word reaction time. The scoring method proposed makes deductions from 100 points for each failure of a subject to meet satisfactorily the four types of requirement just mentioned. Corresponding to the relative importance of the different tests for this purpose a maximum possible deduction of 50 points is allowed for eye movements in test I, a maximum of 25 points each for the distraction and association tests. Stated more in detail, the scheme of deductions for each type of failure is as follows: In the interview test 5 points are deducted for every movement up to ten movements. The maximum deduction is therefore 50 points. For the staring distraction record two points are deducted for each second of time increase in excess of three seconds above his normal adding time. The maximum deduction allowed is 10 points. For the shock distraction 2 points are deducted for every second of time increase in excess of five seconds. The maximum deduction allowed is 10 points. For the snake distraction test 1 point is deducted for each second of time increase in excess of seven seconds. The maximum deduction allowed is 5 points. In the word association test 2 points are deducted for each negative response to 'enterprise,' 'success,' or 'company' and for each vivid, personal type of response to 'death' or 'danger.' One point is deducted for each colorless or doubtful response to 'enterprise,' 'success' or 'company.' The maximum deduction on the basis of the word content of responses is thus 10 points. For the time of response to 'enterprise' and 'danger' 1 point is deducted for every fifth of a second in excess of two seconds. The maximum deduction allowed on this account is 15 points.

The score of any individual subject in aggressiveness is to be obtained by totaling his deductions and subtracting the amount from 100 as shown in table VI. On this basis the scores of our thirteen most aggressive subjects are as follows: C. N. 97, Z. J. 100, A. C. 79, R. M. 94, N. A. 98, J. G. 86, E. B. 83, B. W. 100 (lacking word test), R. W. 93, H. C. 98, A. J. 97, E. L. 84, D. R. 100 (lacking the word test). The scores for the thirteen least aggressive subjects are as follows: H. M. 35, J. J. 79, E. R. 55, W. P. 60, N. F. 59, P. G. 72 (lacking word test), F. B. 36, J. M. 81, G. D. 63, R. B. 37, S. H. 88, R. T. 61, B. R. 38 (lacking word test).

TABLE VI
FINAL SCORES OF 26 SUBJECTS

	Eye Movement	Staring	Shock	Snake	Words	Time	Score
Most Aggressive							
C.N.	..	..	..	2	1	..	97
Z.J.	..	..	..	.	.	..	100
A.C.	10	..	10	.	1	..	79
R.M.	5	..	..	.	1	..	94
N.A.	..	2	..	.	.	..	98
J.G.	5	..	2	4	3	..	86
E.B.	..	6	..	1	4	6	83
B.W.	..	..	..	.	x	x	100
R.W.	5	..	..	.	2	..	93
H.C.	..	..	..	.	2	..	98
A.J.	..	..	..	.	3	..	97
E.L.	5	2	6	.	3	..	84
D.R.	..	..	..	.	x	x	100
Least Aggressive							
H.M.	35	10	10	3	6	1	35
J.J.	5	..	10	4	2	..	79
E.R.	20	4	2	5	4	10	55
W.P.	20	..	10	.	6	4	60
N.F.	20	..	..	.	6	15	59
P.G.	25	..	..	3	x	x	72
F.B.	35	4	..	5	5	15	36
J.N.	..	..	..	.	4	15	81
G.D.	25	..	..	2	5	5	63
R.B.	50	..	2	5	4	2	37
S.H.	..	4	..	.	5	3	88
R.T.	25	4	..	.	3	7	61
B.R.	50	10	..	2	x	x	38

The average score of the aggressive group is 93 and that of the unaggressive, is 58.8. None of the aggressive subjects score less than 79, whereas nine of the least aggressive make less than 64, and one of the others is incomplete. Three of the twenty-six subjects, J. N., S. H., and J. J. appear as noticeable exceptions. It is the opinion of the writers, based on considerable acquaintance with the subjects, that the error in regard to at least two of them lies in original personal ratings rather than in the finding of the test. In any event we seem to be justified in stating that there is hardly one chance in twenty-five that a man weak in aggressiveness would score as high as 85; and there is almost no chance that a very aggressive person will score lower than 70. The writers believe that this test approximates a true measurement of aggressiveness more nearly than does the Army Alpha examination approximate the measurement of intelligence.

Moreover, the test can be given in modified and abbreviated form with no further equipment than a stop watch and a set of addition tables, with the additions all made according to the specifications in test II. In this case it would be necessary to omit the shock and snake distraction tests; also to give the words for the association test orally, and to time the responses with a stop watch. The loss of two of the 10 point tests can be compensated for in the scoring by adding 25% to the total of the deductions made on the basis of the remaining tests.

Aside from the time required of each individual subject for practicing his 50 additions with the tables, the actual time of administering the abbreviated form if the test should not be more than three minute. The longer form with apparatus requires about seven minutes per subject.

THE BEDAUX-UNIT PRINCIPLE OF INDUSTRIAL MEASUREMENT

Its Origin, Its Laws and Applications

By Charles E. Bedaux

(This article is copyrighted by the author, and reproduction in any form is forbidden except by his express permission.)

Ever since industry reached the tremendous proportions afforded by the mechanical era, the need has been felt for a system of measurements permitting the exact determination of the quantitative value of the human physical effort.

It is due to the lack of such a system that to-day we meet deeply set differences between management and labor. Either side ignoring the value of a given effort made it possible for the employee to practice, undetected, a restraint in production, while the employer, in some instances, protected by the same ignorance, has been able to arbitrarily decrease the money value of the operation performed.

A correct system of measurement of the industrial effort has many applications:

1. It permits a protective and adequate system of remuneration;
2. It gauges the capacity of the human being or of a group of employees comprising one department, permits an exact comparison between departments, between plants, and automatically indicates temporary or permanent conditions that are conducive to failure;
3. It may be used as a basis for apportionment of the overhead burden;
4. It allows an automatic and constant comparison between the standard and actual labor cost;
5. It renders scientific planning and scheduling of manufacture possible;
6. It is the only foundation on which any form of labor representation in shop management may stand.

The exponents of the piece-work system have recognized that the basic weakness of their method lies in the inability to guarantee the rate set, as it is based on the monetary value of labor, a commodity which fluctuates. To guarantee the

rate is to expose the manufacturer to failure, in case of a general decrease in the labor market. To cut the rate destroys the confidence of the working man and compels a restraint in production.

The monetary value of labor, being a variable, cannot be used as a means to measure industrial effort.

Because of fluctuations in this value, not only between periods of depression or growth, but also within each period, among the various classes of labor, it has been impossible for the manufacturer to gauge the respective productive capacities of his various departments, or of his entire plant, in comparison with a competitive or co-operative institution.

The result has been that many plants, sincerely believing that their productivity rate was high, have failed because of their error, the detection of which was impossible without an adequate system of measurement of the human effort applied to industry.

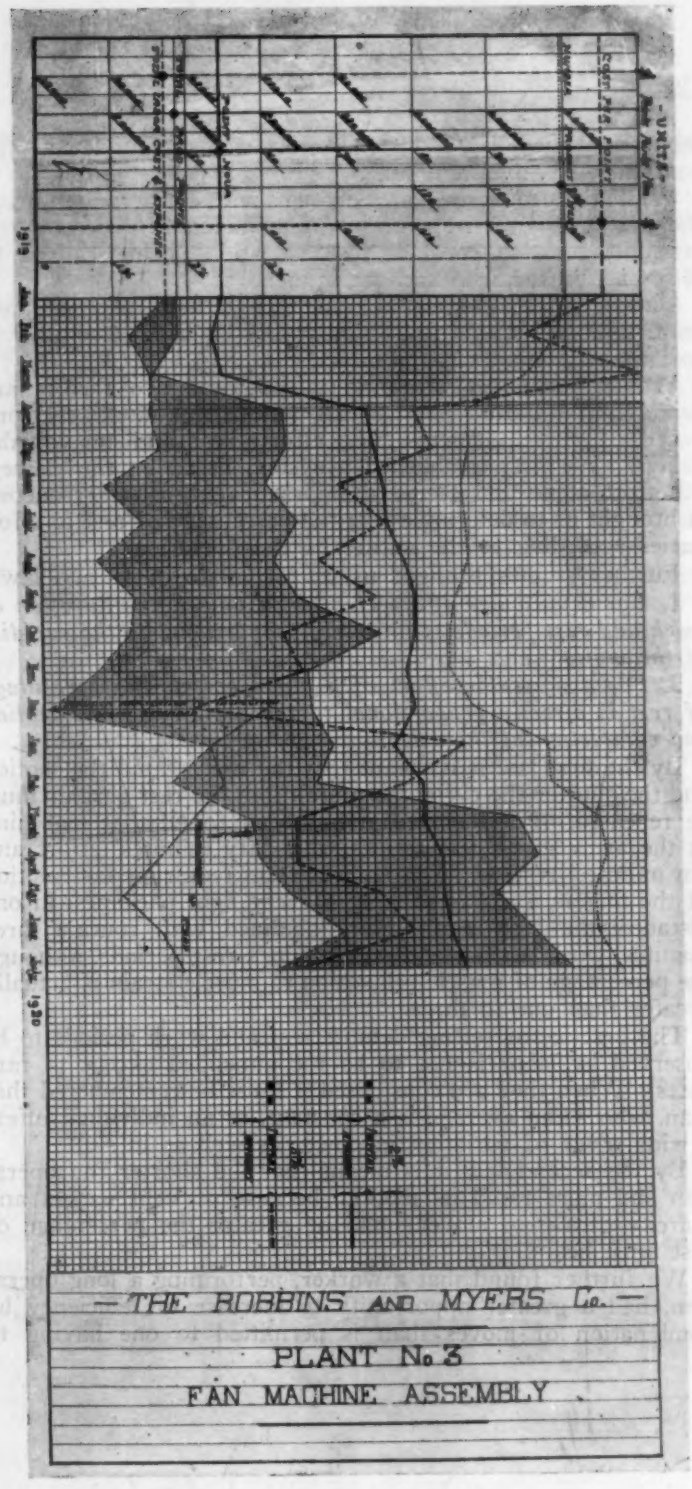
It is also because of the absence of such a system that the monetary value of productive labor has been recorded for cost purposes by the taking of the time of the beginning and the ending of each operation. This performance is generally inaccurate, or, when accurate, necessitates waiting time for recording—a waste of labor, both in the productive and in the clerical forces.

To that source also must be traced the erroneous distribution, too often met, of the overhead burden, wherein an operation costing one dollar ($1.00) in productive labor, performed on a machine worth one thousand dollars ($1,000.00) and depreciating in twenty years, for instance, will be charged with the same amount of overhead as an operation also costing one dollar ($1.00) in productive labor, but performed on a machine the cost of which is one hundred thousands dollars ($100,000.00) with a rate of depreciation that must be absorbed in ten years.

Because of the lack of an adequate system of measurement, incapable operators have been kept in positions that have been refused to men who were best indicated to hold them; promotion has been arbitrary and illogical, thus entailing losses to both capital and labor and creating a state of mind on either side which is wholly destructive.

Part Two

For many years we have seen the need of a unit of measure and our study first led us to determine that such a system could not be based on monetary values which fluctuate. We



— THE ROBBINS AND MYERS Co. —

PLANT No. 3

FAN MACHINE ASSEMBLY

then established that industrial effort was surrounded by two constants, i.e., man and time.

While the productive value changes with every human being, the average power of a large group, taken in different generations, remains constant, the variations being caused by difference in facilities and not in men.

Time is, of course, also a constant. It therefore appeared, on first consideration, that time could be used as a unit of measurement for industrial effort.

We experimented with the use of these two factors and found that the function we call "industrial work" is composed of three elements, i.e., effort, rest and unavoidable delays. We first established a system based on the element of effort alone and found it impractical, because erroneous. It brought us to determine that the percentage of rest to effort varies according to the nature of the operation.

Further studies brought us to formulate these two laws:

1. *For a muscular effort of a given power, the duration of work and rest periods is inversely proportional to the rapidity of the motion,* and

2. *For a muscular effort of a given power, the percentage of rest is directly proportional to the rapidity of the motion and completion of the cycle.*

By the first law it is meant that the more rapid the motion and the shorter the cycle, the more often the rest periods must be repeated. For instance, it has been found that marching in the army gives best results with fifty minutes' effort and ten minutes' rest, while in an operation requiring fast motion of the fingers, such as packing of very light packages, in one instance, the best results were obtained with routine three minutes' work periods and one minute's rest periods, although the percentage of muscle relaxation to muscle strain is actually greater than one to three.

The practice of setting definite rest and work periods to be observed by labor is not to be recommended except in rare cases. When they are considered it must be remembered that man, who as an average is a constant, as an individual offers a wide range of variation in power as well as in skill.

By the second law it is meant that the shorter the operation and more rapid its motion, because of light weight and scarcity of varied movements, the greater the percentage of rest must be.

We further found that a worker, performing a long operation, had a greater opportunity to increase his efficiency by combination of moves than is permitted to one having to

perform an operation the cycle of which is short, therefore demanding a positive routine of a few movements.

On the strength of the above conclusions, our next step led us to compile tables of all ratios of rest and unavoidable delays affecting the various forms of human physical effort in industry. We found it to vary from 15% of the actual productive time in long cycle hand assembly operations to 150% in mechanical operations, hand-fed, of a cycle of two seconds or less.

Experiments, analysis and practical experience born out of various applications of the above brought us to formulate a third law.

3. *In a large group of equally trained workers the average production of the tenth of the total number that is ablest does not exceed the double of the average production of the balance.*

The securing of the above information allowed us to establish the Bedaux Point System of Industrial Measurement. The unit was called a " Point." It has the following definition:

A point is a fraction of a minute of work, plus a fraction of a minute of rest, the aggregate of which is always one minute, but the proportions of which vary according to the nature of the operation.

An employee, man or woman, taking advantage of the normal amount of rest, produces therefore sixty points in one hour, this being the standard. An operator producing more than sixty points per hour is, at the end of the day, entitled to a reward based on the difference between the total points produced and the standard of sixty points per hour.

The operator, by application, can increase his production beyond sixty points per hour without fear of suffering an arbitrary decrease in his reward. The point value of his operation or operations, being based on time in relation to average human physical energy, two elements that are constant, is not, therefore, subject to change. To be logical, the manufacturer must guarantee the point value of any operation, as long as the said operation is not changed in process or material.

The exact knowledge of the number of points produced per hour by each man allows an intelligent system of promotion and the exact knowledge of the number of points produced per hour by a group of men or department permits an adequate remuneration of the leader or foreman. This protective system of remuneration based on the point is known as " The Point System of Graded Remuneration."

Part Three

The logical industrial enterprise is one which manufactures an article or articles in normal demand, made by normal men, performing under normal conditions.

The rest and delays allowances figuring in the establishment of a point representing normal conditions, the aggregate of the elements forming a point representing one minute, the logical manufacture must be profitable at a production of sixty points per hour.

Whenever operators produce more than sixty points per hour, the difference must figure in the element of reward of the operator and of those directly connected with production and whose co-operation labor seeks, with a view to secure the maximum production.

The gain of the manufacturer is realized in the reduction of the overhead burden per unit.

The industrial enterprise which finds any or all of its departments producing daily at a rate of less than sixty points per hour knows it is in danger and can therefore take immediate steps to bring the point hour to sixty or higher in any or all of its departments.

Manufacturers having branch plants can compare the point hour of each plant each day and thus measure the advance of the various units of their enterprise.

The aggregate point value of an article represents the standard labor cost. The accumulation of all the points charged to a given factory order allows automatic comparison of the actual cost against standard cost, this without the expensive recording of the time of beginning and ending of each operation.

The unit of measurement of the Bedaux system, the point, can be used in the distribution of overhead burden by determining in advance the cost of running one point through each manufacturing unit and multiplying this cost by the point value of all operations passing through, thus giving an exact overhead distribution.

Countless efforts have been made in the past to plan, schedule and dispatch material through manufacture. Many failures have resulted, for it must be remembered that planning is a relation of requirements to capacity.

To determine the time requirements of an article passed through different branches of manufacture, or to determine the time capacity of the various branches of manufacture through which the article is to be passed, one must have at

his disposal a unit of measure, which, in industrial effort, has heretofore been unknown.

With the Point, planning becomes an easy matter. The point value is known on every article in all its operations. The point capacity of all machines or branches of manufacture is known through the point hour. The work of planning thus becomes simply a routine detail relation of point value to point hour.

Part Four

Plans of profit-sharing have been established in the past with a questionable degree of success, only because there has been no way to measure the capacity of each individual when determining the share of the profit to which he was entitled. If the manufacturer sincerely desires to make his labor share in the profits, the amounts can be apportioned according to the point hour of each man in relation to his class of skill, which is gauged by his hourly rate.

Various forms of labor representation in shop management have lately been widely praised as the logical panacea for industrial ailments, but is it logical to organize a labor representation which is formed indifferently by men who are capable or incapable? Has it not always been the rule for man to master the capacity to perform before he assumes the capacity to advise? The point hour record of each man in his daily labor together with his mental capacity should alone determine his eligibility to representation of his fellow workers, whenever the manufacturer feels that he must adopt that form of management.

The most formidable obstacle to overcome in establishing industrial peace is the present restraint of hope caused by the industrial structure of to-day. The bench man remains a bench man. A press operator, an inspector, an automatic screw man, a lathe operator, all remain what they are. It has been so customary for industry to endeavor to keep a man wherever he is doing "fairly well" that labor itself has lost the thought of possibility of individual rise but in turn has reached a state of mind generally indifferent or else bitter and antagonistic.

Under the Point System of Industrial Measurement, an unskilled laborer who, by application, shows a point hour of, say, ninety, or thirty points above standard, this for a period of a year, is automatically entitled to a better position whenever there is a vacancy and he can begin such at a point hour of not less than sixty. If, by further application, he brings

himself, in this new and better position, to a point hour of ninety and maintains himself there, again he is automatically entitled to a still better position, at a higher hourly rate, providing he can begin his performance, through previous study, at a point hour of not less than sixty. This can be continued through all the branches to the very head, so that when a man is given the direction of a department because he has shown fitness in the elemental requirements of the executive position, it will be known that he has also proven himself capable in all the branches he commands.

At the present time the labor unions, knowing that the chances for advancement in productive labor are small because of industrial tradition, have endeavored to raise the scale of wages of the whole strata composed of the various classes of effort, a suicidal move, as if applied it only raises the price of the commodity purchased by the consuming class, which is largely composed of the class of people to which union labor belongs.

As illogical as this labor demand may seem, there is no logical answer to give other than to render possible the advancement of the individual through each class, according to the exact individual capacity, a capacity which to-day can be accurately measured by the Point without increasing but even decreasing the amount of clerical work usually found in most plants.

A STUDY IN INDUSTRIAL PSYCHOLOGY—TESTS FOR SPECIAL ABILITIES

By Elsie Oschrin Bregman

The following is the report of an investigation with the purpose of developing tests for the use of the Employment Department of R. H. Macy & Co., Inc., a large department store in New York City.

The problem resolved itself into a twofold one; first to determine whether or not the applicant who applies is of normal intelligence, and secondly, to determine, of the normal applicants, the specific kind of work to which they shall be appointed.

The first problem has been handled in a fairly adequate manner by an adapted form of the Trabue Completion Test. The results of this test will, it is hoped be presented very soon in a separate paper. This report concerns itself with the development of tests for specific abilities.

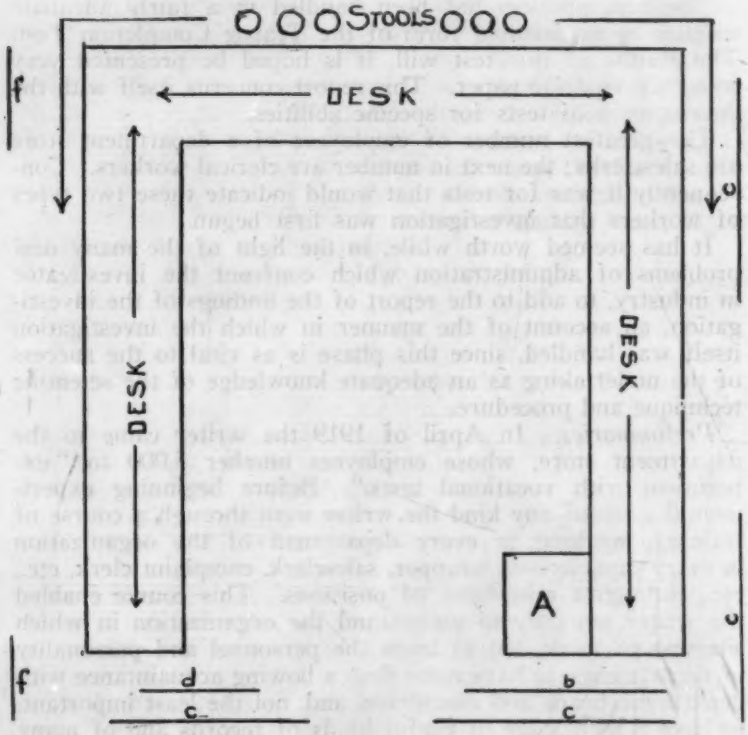
The greatest number of employees of a department store are salesclerks; the next in number are clerical workers. Consequently it was for tests that would indicate these two types of workers that investigation was first begun.

It has seemed worth while, in the light of the many new problems of administration which confront the investigator in industry, to add to the report of the findings of the investigation, an account of the manner in which the investigation itself was handled, since this phase is as vital to the success of the undertaking as an adequate knowledge of the scientific technique and procedure.

Preliminaries. In April of 1919 the writer came to the department store, whose employees number 5,000 to "experiment with vocational tests." Before beginning experimental work of any kind the writer went through a course of training, working in every department of the organization in every capacity—as wrapper, salesclerk, complaint clerk, etc., etc., through a miscellany of positions. This course enabled the writer not only to understand the organization in which she was to work, but to learn the personnel and personality of departments, to have more than a bowing acquaintance with department heads and executives, and, not the least important, to have a knowledge of useful kinds of records and of many sources of information which blueprints and verbal reviews did not give.

Immediately following this course of sprouts, the writer became an active member of the Employment Department, interviewed and appointed applicants and during three months (during which time an office which could be used as a laboratory was being built and equipped, and preliminaries to actual experimental testing, such as collection and printing of forms, etc., were carried on) acquired a grasp of the local labor situation, the sources of supply, the types, characteristics and social levels of applicants who presented themselves, the ways in which appointments and judgments were made, etc., etc.

The following diagram illustrates the office laboratory which was built where the testing could be carried on with the minimum of noise and disturbance.



a. Examiner's Desk
b. Bookcase
c. Windows
d. Blackboard
e. Cupboard
f. Doors

Before the testing itself took place, meetings of the executives, heads and assistant heads were called, at which the purpose, plans and method of the investigation were disclosed and explained as simply as possible, and cooperation was urged. In order to dispel prevalent bugaboos about the nature of mental tests, a five minute Trabue Completion Test was given.

Obtaining the Criteria. In view of the unfortunate waste that has not infrequently occurred when after extensive tests had been made, records of work and ability were unobtainable, it seemed wise to secure the criteria before testing, the more so that some coherent scheme of rating had to be worked out in an organization where for two thousand salesclerks there were as many as 200 department and section heads.

Salesclerks. Considering first the salesclerks, two methods of determining their ability were available (1) the actual weekly sales or production record made by each clerk, (2) the opinions of department heads, of which there are for each department

1. Buyer
2. Assistant Buyer
3. Section Manager or Floorman.

Production Record. Since the gross amount of sales varies for each department and from week to week, according to season and weather, it was not possible to throw all sales on to one scale for all time. The following method was used.

A simple frequency table was made showing the sales for each department for each week. From this the upper and lower quartile limits were determined, and record was made for each clerk, whether he sold within the interquartile range, above it, or below it. Such a record was kept for ten weeks and from this was determined whether a clerk sold characteristically within, above or below the average range.

Ratings. This information was supplemented by the ratings of department heads, which were obtained in the following manner:

The accompanying form was sent to each buyer and assistant buyer and section manager. From the information so gained it became possible to group the bulk of the selling force into three large categories—the good, the poor and the average salesclerks.

Sec. Man.
To M Buyer Dept.

Please send to ..... of the Employment Office, on this form, not later than....., 19..., three lists of

employees, carefully made out according to the following instructions.

I. List below the names of as many salesclerks as you know who are now employees of R. H. Macy & Co., Inc., and who represent, in your opinion, the most capable and desirable type of salesclerk.

Write down **only** the names of **absolutely first class** salesclerks; those who really are able to sell.

No.	NAME	Remarks

II. List below the names of as many salesclerks as you know who are exactly the opposite of the people you have just named, representing the poorest type of salesclerk with little or no selling ability.

No.	NAME	Remarks

III. Place below a list of salesclerks who are, in your opinion, neither very good nor very poor; the ordinary type of salesclerk who is satisfactory but not exceptional.

No.	NAME	Remarks

This loose form of rating sheet was used in preference to a more rigid and detailed form for the following reasons:

(1) The department heads of the organization under consideration are to a great degree free agents. It was essential for the return of the largest number of these sheets with a minimum amount of energy and friction entailed, that the form be filled out with the utmost ease.

(2) The department heads are not an academic group—the median intelligence lower than that of such a group—they had consistently shown an aversion to any sort of analytic thinking or patient arrangement of groups into rank order, and such thinking, if forced upon them would be of a doubtful reliability. It seemed on the whole to be wisest to use a spontaneous method of rating—the more so that it would have been impractical to test the entire force of salesclerks. Therefore, by this method of rating, the testing was confined to those individuals who were spontaneously recalled in the minds of the several department heads as meriting good, bad or indifferent ratings.

The results of this method bore out the choice of the course. The forms were returned on time and with a minimum of follow up work. The large body of salesclerks was rated and the undertaking did not lose in prestige by being called superscientific, impractical, etc.

The judgments so obtained from these three sources, buyer, assistant buyer and section manager were then matched up on the following form.

Department.....

Disc. No.	NAME	Sales	Buyer	Ass't Buyer	Section Manager	Remarks

Fifth, miscellaneous groups of individuals about whom there was no common and consistent type of judgment—those rated as both good and poor, for instance. These were wholly omitted from the experiment because of the inconsistency with which they were rated.

Clerical. The other large group of employees engaged in work of generally similar character were the clerical workers. For the same reasons and in the same fashion as applied to salesclerks ratings of the clerical workers were obtained previous to testing for the three large clerical departments—the Audit, the Mail Order and the Complaint Department. Unlike the salesclerks, however, there was no common production record by which the judgments of department heads could be controlled, the clerks being engaged in a multiplicity of operations, from complicated statistical work and bookkeeping, to counting of checks, filing and comptometer operating. Moreover in most cases the clerks were well known to only one person, the local supervisor, only superficially known to the actual department head and very occasionally known to as many as three persons.

Raters were therefore instructed to rate only those clerks whom they knew and the rates taken at their face value, with the precaution that such clerks about whom there was a wide divergence of opinion were excluded from the possible subjects for experiment, and only such clerks included about whom at least two people were agreed.

The following scale indicates the manner in which the clerks were grouped.

1. Rated as first class by two or more—no dissenting opinion.
2. Rated as first by two—2nd by 1.
3. Rated as first by one—2nd by 1.
4. Rated as second by two—1st by 1.
5. Rated as second by two or three—no dissenting opinion.
6. Rated as third by one—2nd by 2.
7. Rated as three by one and 2nd by 1.
8. Rated as three by two and 2nd by 1.
9. Rated as three by three or two—no dissenting opinion.

In obtaining ratings throughout the course of this work, it was noted that there was a very marked reluctance to say that an employee was third rate or unsatisfactory, for the reason that the executive giving such a rating laid himself open to the question of why he retained such a person. It seemed fair, therefore, to infer that all clerks whose total rating was more than seven, could be included among the

third rate clerks, and they were so considered. This was also done because the number nine group alone made up a very inconsiderable number. The ratings of *good* clerks, if unanimous, were significant; of the poor and mediocre clerks, less so because of the hesitancy to call any one poor, and of a tendency to underrate if the clerk was unknown or doing unimportant work, or if but newly employed.

However, these ratings, such as they were, served to indicate three groups, the very good (1) the poor (7-9) and the generally average clerks who were rated between these two extremes.

Tests. Through the courtesy of Prof. E. L. Thorndike and the National Research Council Committee sets of the N. R. C. Tests, the precursors of the National Intelligence Tests were made available for the purposes of this investigation.

These tests were:

1. Verbal A which consisted of
   1. Arithmetic
   2. Directions
   3. Sentences
   4. Synonym Antonym
   5. Judgment
2. Non Verbal A
   1. Picture Completion
   2. Series Completion
   3. Comparison
   4. Symbol Digit
3. Verbal B
   1. Computation
   2. Vocabulary
   3. Sentence Completion
   4. Disarranged Sentences
   5. Logical Selection
4. Non Verbal B
   1. Copying Designs
   2. Pictorial Sequence
   3. Pictorial Identities
   4. Recognitive Memories

In addition the following tests were also used,

[1] Sentence Completion 10′ and 5′ time limits—called T1A

Arithmetic—Woody McCall Mixed Fundamentals—called T2

Rearrangement of Animals—An abbreviated form of the writer's test published in June, 1918, *Journal of Applied Psychology*—called T3

Woodworth Wells Mixed Relations

Woodworth Wells Mixed Relations changed to an underlining test—called T5

Woodworth Wells Opposites

Woodworth Wells Opposites changed to an underlining test—called T4

Woodworth Wells—Number group checking

Woodworth Wells—Cancellation

[1] This test is a scale of 25 Trabue Sentences, arranged by the writer, for which norms from approximately 20,000 adults are at hand.

Testing. Directions for performing the tests were identical for every individual.

The tests were given to groups in the laboratory sketched above. The testing was always in the early morning—from 9 to 10:30 A. M. Meetings did not last longer than 90 minutes. Subjects were recalled for a second period if the tests were not completed in the first.

Before beginning any testing, the purpose and method of the work was carefully explained to each group as simply as possible. Discussion was perfectly free and testing was never begun without good feeling and rapport between the subjects and the experimenter. Pains were taken to dispel any preliminary nervousness, and to see that subjects took the tests comfortably. Both sexes and all ages from 17 to 60 were represented. Of the salesclerks there were sellers of every type of merchandise from upholstery and women's suits to shoes, veiling, gloves and notions. Every phase of clerical work at which the employees of this establishment were engaged was likewise represented in the test groups. There were skilled typists and stenographers, statistical workers, bookkeepers, billing clerks, comptometer operators, check counters, filers, keepers of simple records, etc., etc.

For the initial phase of the experiment, in which a rough evaluation of the tests as to their sensitivity to different degrees of abilities was sought, only the extremes of both groups were used; that is, of the salesclerks, those were tested who were **consistently** classed as good and as poor; of the clerical workers, those who, according to the classification scheme indicated above, were classified as 1, and those who fell within the 7 to 9 classifications.

The average workers of each group were made use of in a later phase of the experiment.

Correlation Formula and Statistical Methods

Several correlation methods were used in the course of this work.

1. The method of Unlike Signed Pairs—discussed in Thorndike's Mental and Social Measurements, p. 162, pp. 170-1—and called in tabulations and discussions in this paper the U Formula.

2. Pearson Biserial R.—a formula which determines the correlation when one variable is measured and continuous, the other unmeasured and alternative—and which closely approximates the $r = \frac{\Sigma xy}{n\sigma x\sigma y}$ see *Biometrika*, Vol. VII, 1909, p. 96.

3. Finally the orthodox Pearson $r = \frac{\Sigma xy}{n\sigma x\sigma y}$ was used in

the last phase of the experiment, in partial correlation and regression in order to weight the significant tests. When this formula was used the value 1 was given to a quality i.e., good salesmanship, and 0 to the absence of it.[2]

The other formulas do not necessitate so arbitrary an assignment of values.

Correlation Coefficients—Salesclerks. Table I below represents the coefficients which were obtained from the salesclerks when the good and the poor clerks were tested with a series of tests. The number of clerks tested, the names of the tests and the statistical method used to obtain the coefficients are all indicated.

TABLE I

GOOD AND POOR SALESCLERKS

Number of Cases	Tests	r by U Formula
45	N R C – N V A Total	— .09
46	" " " " " " 1 Picture Completion	— .39
44	" " " " " " 2 Series Completion	+ .22
44	" " " " " " 3 Comparison 3′	— .59
44	" " " " " " 3 Comparison 2′	— .83
43	" " " " " " 4 Symbol Digit	+ .25
44	" " " " " " 5 Form Combination	— .37
44	" " " V A Total	— .59
43	" " " " " 1 Math.	— .16
44	" " " " " 2 Directions	— .66
43	" " " " " 3 Sentences	— .79
44	" " " " " 4 Synonym Antonym	— .22
45	" " " " " 5 Common Sense	— .79
43	" " " " " 6 Analogies	— .54
46	" " " N V B Total	+ .16
48	" " " " " " 1 Copying Designs	— .06
46	" " " " " " 2 Pictorial Sequence	+ .16
48	" " " " " " 3 Pictorial Identities	.00
45	" " " " " " 4 Recognitive Memories	+ .51
48	" " " V B Total	— .06
49	" " " " " 1 Computation	— .16
49	" " " " " 2 Vocabulary	.00
49	" " " " " 3 Sentence Completion	— .34
49	" " " " " 4 Disarranged Sentences	— .28
49	" " " " " 5 Logical Selection	.00
48	T 1 A Sentence Completion 10′	— .13
48	T 1 A Sentence Completion 5′	— .37
33	T 2 Arith.	— .16
49	W. W. Opposites	— .16
47	T 4 W. W. Opp. Adapted	— .43
48	W. W. Mixed Relations	— .37
48	T 5 W. W. M. R. Adapted	— .46
48	T 3 Rear. Animals	— .56

[2] This method of statistical treatment was suggested by Prof. T. L. Kelly, now at Leland Stanford University.

Some of these coefficients are markedly high, also they are, almost without exception, negative.

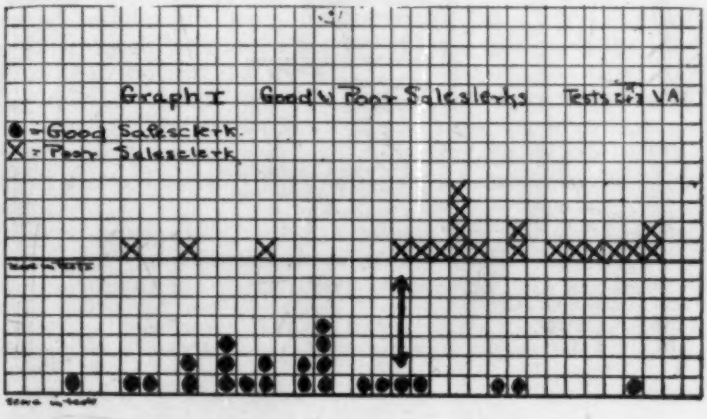
Table II presents coefficients which were obtained from combining in various groups the tests which singly gave the highest coefficients. (The coefficient for NVA3—2′ time limit had not been computed at the time these groupings were made and the coefficients of Table II calculated.)

The general tendency, it will be noticed, is to increase the correlation coefficient by such combinations of tests.

TABLE II

GOOD AND POOR SALESCLERKS' TESTS COMBINED

Number of Cases	Test	r by U Formula
41	Sum V A 3 N VA 3 — 3′	— .66
39	" " " 2-3-5-6 N V A 3 — 3′	— .69
43	" " " 3-5	— .79
43	" " " 2 and 3	— .84
44	" " " 2 and 5	— .78
42	" " " 2 and 6	— .56
44	" " " 2 and N V A 3 — 3′	— .66
42	" " " 3 and 6	— .64
43	" " " 2-3-5	— .84
42	" " " 5 and 6	— .51
44	" " " 5 and N V A 3 — 3′	— .61
39	" " " 6 and N V A 3 — 3′	— .75
42	" " " 2-3-6	— .79
40	" " " 2-3 N V A 3 — 3′	— .66
42	" " " 2-5-6	— .68
40	" " " 2-5 N V A 3 — 3′	— .66



Graphically represented these coefficients take on a more concrete meaning.

Let us consider for a moment the practical implication of such a situation as Graph I represents. Suppose that the forty odd people represented by the sum of the crosses and the rows of dots were, without any distinguishing mark, seated in a room. By means of tests which are represented in Graph I (using the position of the arrow as a reference point) an individual would be able to determine which of the group had been successful and which unsuccessful as salesclerks, and would be in error only 8 times out of the forty-four cases.

Clerical. The clerical workers, good and poor, represented by classifications 1 and 7-9 were tested with the following tests,

V A—All
N V A—All
W. W.—Cancellation
W. W.—No. gr. checking
Rearrangement of Animals—T3

Table III which follows presents the coefficients which these tests gave, the formula in use being again that of unlike signed pairs.

TABLE III

GOOD AND POOR CLERICALS

Number of Cases	Test	r by U Formula
59	N R C – V A	.34
61	" " " " " 1 Arithmetic	.09
61	" " " " " 2 Directions	.00
60	" " " " " 3 Sentences	.31
58	" " " " " 4 Synonym Antonym	.28
60	" " " " " 5 Judgment	.37
50	" " " " " 6 Analogies	.28
48	" " " N V A	.31
52	" " " " " " 1 Picture Completion	.13
52	" " " " " " 2 Series Completion	.22
51	" " " " " " 3 Comparison 2′	.45
50	" " " " " " 4 Symbol Digit	.37
52	" " " " " " 5 Form Combination	.00
34	No. Gr. Ch.	.31
35	Cancel.	— .13
57	T 3	.19

It is evident at once that the ratios are not as high as those obtained for the salesclerks, the highest being .45 for N V A 3—the Comparison Test.[3] However, it is also immediately evident that these ratios are all positive, with the one exception of Cancellation. Now it will be remembered that the N V A and V A tests were the tests which gave the highest correlations for salesclerks but that these ratios were almost all consistently *negative*.

It seems fair to assume that had it been possible to secure clerical workers rated with the freedom from error of the salesclerks' classifications, and in view of the fact that a positive relationship is indicated by the plus ratios, that positive correlations perceptibly higher than those actually found would have been obtained. However, this is merely an assumption; what is apparent is the tendency for salesclerks and clericals to pull in distinctly opposite directions, so that we have a difference between trades as distinct as the differences within a trade, if not more so.

This is a not unimportant consideration, for the primary interest of an Employment Office is not in deciding with what degrees of excellency an unknown individual will perform a given task, but to decide at what task to place a given individual.

Concretely, Mary Jones, age 18, without experience or training wants a job. The Employment Manager wants to know whether to make Mary Jones a salesclerk or a clerical. To him whether Mary Jones will rank fifth or first as a salesclerk is a matter of academic interest only. How good she will be is a matter she must show by performance, since in industry as elsewhere reward comes not in anticipation of work but follows it. Wages in a large industry are fairly standardized at employment time, and are modified not for potential ability but only for previous experience. They become really differentiated for individuals only with the history of that individual's work.

In order to express numerically the contrary tendencies of salesclerks and clerical workers, the two groups were thrown together and an attempt was made to separate from this heterogeneous group those who were called good salesclerks and good clericals.

Such a combination seemed to have two virtues, (1) the reliability of the coefficients obtained would be increased because the size of the group would be doubled; (2) the hetero-

[3] It will be remembered that this test gave the highest correlation coefficient, but a negative one, when used with salesclerks.

genity of the enlarged group more nearly approximated actual working conditions. Candidates for employment as they present themselves at the office which these tests were to serve are not divided off into sales or clerical workers and often have no strong preference for any special kind of work, but must be classified and selected for various jobs by the individuals who interview them.

This combination of both types of workers was made in two ways, first all sales and all clericals were combined, and selection made, for (1) good salesclerks, (2) good clerical workers. The second grouping was as follows. When good salesclerks were being selected the unsatisfactory clerical workers were omitted from the total group of sales and clericals for the reason that there did not seem to be any ground to consider them as undesirable candidates for selling.

All that was known was that they were undesirable for clerical work. Of course it may be argued that the retention of the good clerical workers was likewise ungrounded; that there was no reason why they should not make good salesclerks, even if they were good clerical workers. This is doubtless true. At the same time, considering again the practical aspect, they were actually non-sellers, the negative-positive correlation findings seemed to indicate that also potentially they might be non-sellers, and it seemed to be of interest to see what would happen to the coefficients already obtained when such a group as this under discussion was made.

When the selection for good clericals was made, the unsatisfactory salesclerks were omitted for the same reasons given for omitting unsatisfactory clerical workers when selecting for salesclerks.

Tables IV and V present the coefficients which were obtained from this double grouping. The coefficients which were obtained from the original pure sales and pure clerical groups are also given for comparison, so it may be seen how the coefficients are affected by the increased number of subjects and the different kinds of groups.

Apparently, the elimination of the bad clericals, in the case of the selection for salesclerks, was justified, for when good salesclerks were selected from the heterogeneous group of all salesclerks and all clericals correlations were lower than those obtained with the salesclerks only. However when the poor clericals are omitted from this heterogeneous group and selection then made for good salesclerks the situation appears to have cleared again, for the coefficients regain their original character occasionally being slightly higher or slightly lower

TABLE IV

SELECTING GOOD SALESCLERKS

Test	Number	All Sales and All Clericals r by U Formula	All Sales and Good Clericals only r by U Formula	Number	Good and Bad Sales Only r by U Formula	Number
N R C – V A	102	—.31	—.56	71	—.59	44
" " " " " 1	104	—.22	—.28	75	—.16	43
" " " " " 2	105	—.43	—.56	76	—.66	44
" " " " " 3	103	—.59	—.75	74	—.79	43
" " " " " 4	102	—.12	—.22	73	—.22	44
" " " " " 5	105	—.48	—.64	74	—.79	45
" " " " " 6	103	—.56	—.69	74	—.54	43
" " " N V A 3-2′.....	95	—.96	—.93	66	—.83	44
Rear An.–T3	106	—.19	—.31	80	— 56	48
V A 3–5 N V A 3–2′	93	—.82	—.91	65	.....	..
V A 3–5–6 N V A 3-2′	92	—.92	—.94	64	.....	..
V A 2–3–5–6 N V A 3-2′.....	92	—.69	—.90	64	.....	..

TABLE V
SELECTING GOOD CLERICALS

Test	All Clericals And All Sales		All Clericals and Good Sales Only		Good and Bad Clericals only	
	Number	r by U Formula	r by U Formula	Number	r by U Formula	Number
N R C – V A	102	+.28	+.45	82	+.34	58
" " " " " 1	104	+.10	+.16	85	+.09	61
" " " " " 2	105	+.03	+.19	85	+.00	61
" " " " " 3	103	+.31	+.54	84	+.31	60
" " " " " 4	102	+.18	+.27	82	+.28	58
" " " " " 5	105	+.54	+.61	86	+.37	60
" " " " " 6	103	+.31	+.42	84	+.28	60
N R C – N V A 3-2'	95	+.66	+.64	74	+.45	51
Rear An.–T3	106	+.16	+.34	81	+.19	58
V A 3-5 N V A 3-2'	93	+.54	+.71	76	.....	54
V A 3-5-6 N V A 3-2'	92	+.56	+.79	75	.....	..
V A 2-3-5-6 N V A 3-2'	92	+.54	+.77	75	.....	..

than in the original classification. Their significance is however increased, since their P. E.'s are decreased, because of the increased size of the group from which they were obtained.

When good clericals were selected the combining of all salesclerks and all clericals does not seem to affect the coefficients markedly, except to raise them in several instances. With the clericals, it will be remembered, the waters were originally muddy, so that an addition of several more misplacements would probably not have the same disturbing effect as with the salesclerks, where the original groupings were markedly clear.

However, when from the total group the unsatisfactory salesclerks are withdrawn and selection then made for good clerical workers, the coefficients are, with the exception of one case where the ratio remains practically the same, raised to a marked degree, and again we seem to have a justification for having excluded the poor salesclerks from this group.

Not only are the coefficients raised but in some cases they become large enough to promise a fairly satisfactory basis for the selection of clerical workers, something which the size of the coefficients obtained with the original group of clericals only did not warrant.

Final Phase of the Experiment. Having, in the manner above described, found several tests which seemed sensitive to selling and clerical ability, it became necessary, in order to employ these tests for the guidance of the Employment Office, to evaluate each test, in accordance with its importance in the group of tests. In other words, it was necessary to determine the regression equation for the whole group of tests, and the best weighting for each test in that group.

Most of the tests however which gave significant coefficients were in the group of the National Research Council Tests. These tests had been constructed for use in schools and their content was not especially adapted for industrial use.

People who are looking for a job usually object to being asked who Black Beauty or Thomas Jefferson are and are apt to ask in response "Is this a school we are in?" Even if this response is not audible, such questions do not result in a very favorable state of mind, and since the success of industrial tests is as dependent upon their reception by the people who take the tests as upon the prophetic accuracy of the tests themselves, new tests were constructed, in form

the same as certain of the N R C Tests, in substance different. The atmosphere savored more of calico, pins and sealing wax and less of George Washington, Abraham Lincoln and the states of the Atlantic Seaboard.

The test questions were all placed on one side of a sheet of 8½x11 paper, on the reverse side of which was a dotted line for the name, and several sample questions of the test itself, the samples being of course so simple as to be self-explanatory. The questions on the test paper were arranged roughly in an order of increasing difficulty, the earliest questions being scarcely more difficult than the sample questions.

A term used by Link, "shock absorber," expresses a perfectly sound principle. Non-academic groups have a very strong aversion not only to academic problems, but mental gymnastics of any kind, and it is much better for the prestige of the test to get such people started on primer tasks and then imperceptibly involved in the more difficult tasks, than immediately to frighten and make them hostile with the more difficult questions.

It was necessary, of course, after the new tests had been made to ascertain if they would act in the same way as the tests they were meant to replace, that is, give the same characteristic correlation coefficients.

This called for more testing. It will be recalled that in originally getting classifications for the salesclerks and clerical workers there were several groups, the very good, the very poor, the average, and the slightly better than average. Of these groups the extremes, good and poor of both workers had been tested in the original phase of the experiment. The clericals and salesclerks who were rated as average had not been tested.

It seemed unwise to test the original groups with the new tests, because, having spent some three hours in test taking they had had an amount of practice which made them no longer naïve, and also because it was inadvisable to take those who had already lost a considerable period of working time away from their departments for a further period.

Consequently the new tests were standardized on the group of untested workers, those who made records of average clerks, or in some cases somewhat better, but who were neither absolutely first rate, nor third rate workers. One hundred salesclerks were tested and 43 clerical workers.

The new tests were modifications of the N R C Common Sense—called in the Tables T 9, N R C Directions—T 8, N R C Comparison—T 13, N R C Sentences—T 12. In addition Completion Test T 1 A—10′ time limit—was used, the adapted form of Woodworth Wells Mixed Relations—T 5, and the Animal Rearrangement Test—T 3.

The formula which was first used to evaluate these tests was the Pearson biserial r, the clerical workers,[4] 43 in number being considered the special group and the 100 salesclerks being considered non-clerical workers.

The coefficients obtained with these tests and this group of workers is given below in Table VI.

TABLE VI

	Test	Pearson Biserial r	Number
T 1 A	Completion	.19	143
T 8	Directions	.63	143
T 9	Sentences	.45	143
T 12	Judgment	.55	143
T 13	Comparison	.74	143
T 3	Rearrangement of Animals	.11	143
T 5	Mixed Relations	.51	143

With the exception of T 3 and T 1 A the coefficients are high and the performance of the two groups, sales and clerical, is consistent with the performance of the groups previously tested, the salesclerks giving low scores, the clericals high scores.

T 3 and T 1 A could therefore be omitted from any final group of tests. It was decided, in spite of the size of the coefficient of Test 5, which is W. W. Mixed Relations, to leave it out of consideration also, because of the difficulty always encountered in explaining the task and having it understood by those tested. With the other tests it was simply necessary for the subjects to see and do the sample problems, without other explanations from the experimenter. For the sake of economy in time and effort in the administration of the tests, when they should be used as part of the employment process, it seemed wise to discard this test, which had to be explained to the majority of subjects.

[4] It was at the time impossible to test more clerical workers because of the Christmas season rush.

The next step was to find the regression equation for the group of tests which were retained and the best weighting for each test.

The formulae in partial correlation and tests weightings are derived from the Pearson coefficient $\frac{\Sigma xy}{n\sigma x\sigma y}$. The coefficients in Table VI are those of the Pearson biserial r, which closely approximates this r, but which is not identical with it.

As stated previously, the $\frac{\Sigma xy}{n\sigma x\sigma y}$ values for the new tests were found by giving the presence of sales ability the value 1, the absence of it or as was actually the case, the presence of clerical ability the value 0, the test scores forming the values of the second variable.

The regression equation was thus determined for 4 variables and the criterion, the variables being the scores in the Directions, Comparison, Judgment and Sentence Tests. The weighting of the Sentence Test however was so very inconsiderable that it was deemed inadvisable to retain it in the group and consequently a second regression equation was calculated in which there were only 3 variables and the criterion.

The following regression equation was obtained using the variables:

X1 The criterion —
X2 Comparison Test —
X3 Direction Test —
X4 Judgment Test —

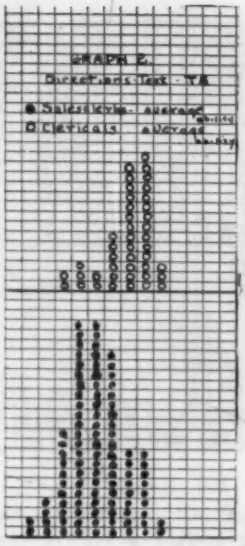
$X_2 = .025\ X_3 + .065\ X_3 - .007\ X_4 - .85$[5]

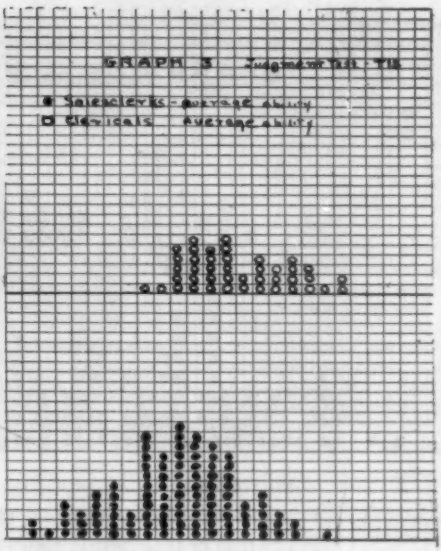
R for this equation = .63
Biserial r = .74

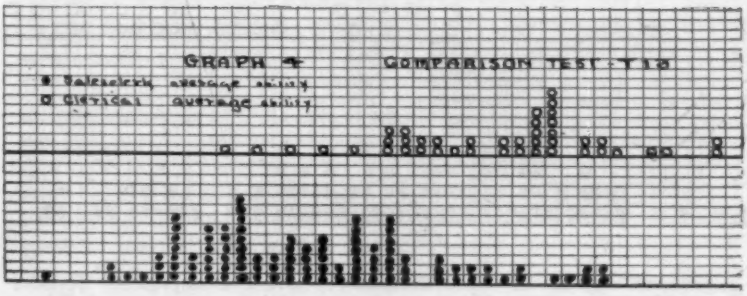
The manner in which, by weighting the tests in accordance with the regression equation above stated the division between the two groups, sales and clericals is emphasized is indicated by the following graphs.

Graphs 2, 3, 4 represent the performance of the sales and clerical groups in single tests—T 8, T 12, T 13. The tendency for clericals to do better work is evident to some degree. A great number of misplacements would occur however if estimates of ability were made by these graphs as they stand.

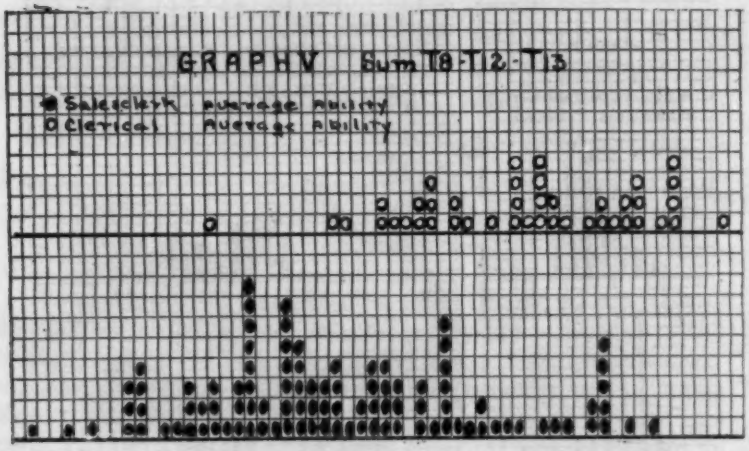
[5] This equation is expressed in gross scores.

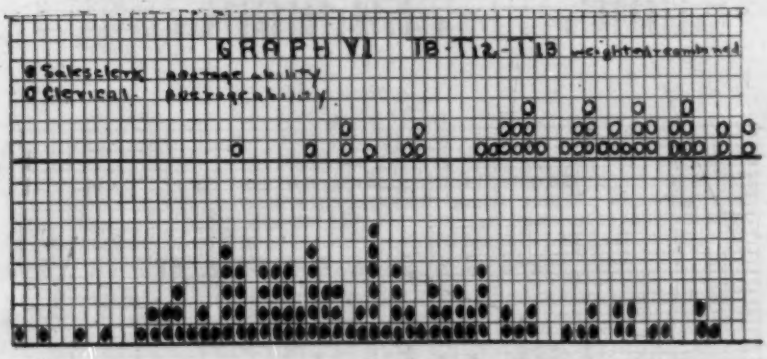

GRAPH 2.
Pocket-Aris Test - T8
● Salesclerks - average ability
□ Clericals average ability


GRAPH 3
● Salesclerks - average ability
□ Clericals Average ability


GRAPH 4
COMPARISON TEST - T 10
● Salesclerks average ability
□ Clericals average ability

Graph V is the sum of these tests. Here the two groups show a marked tendency to separate, but there is considerable overlapping. Graph VI represents the tests weighed and combined in accordance with the regression equation.

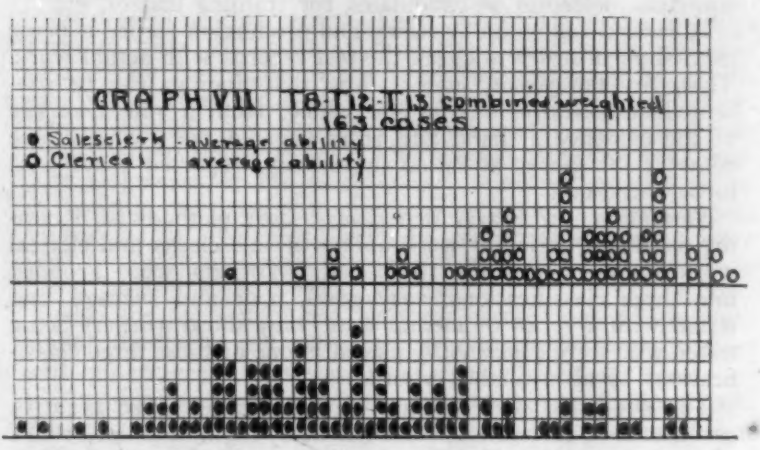




The overlapping has been diminished—and the salesclerks form a fairly clear characteristic group, the clericals as distinctive a group.

After these regressions and other calculations had been made for the 43 clerical workers, it became possible to test 20 more clerical workers. It seemed inadvisable to recalculate on the basis of the enlarged group the weightings as originally found, but it did seem of interest to see what effect adding these 20 cases would have on the coefficient, when the tests were weighted in accordance with the values that had been obtained from the smaller group.

Pearson biserial r was raised from .74 with 143 cases to .83 with 163 cases. Graph VII shows the increased emphasis on the separation of the two groups.



Applications. All applicants for employment, during the time in which the experimental work above described was being carried on, were being tested with the Completion Scale which has been previously mentioned, and which served to indicate mental defectives and subnormal individuals. Upon the findings discussed above the group of tests T 8, T 12 and T 13 were added to the test already in use and all applicants who were considered for either selling or clerical workers were tested with this group, and placement made as far as possible in accordance with the characteristic range into which the test scores fell, whether sales or clerical. The Completion Scale was still retained and in every case given to the applicant before the group of tests. Thus the fact that a low score, or one in the range characteristic of salesclerks was due to general mental inferiority, and not other causes was guarded against.

The tests have been in use for some time. In about 50% of the cases interviewers have applicants tested before placement, the other 50% of the cases are placed and then tested—the placement being subject to revision when the test findings strongly contradict the placement.

Applicants are tested generally by a clerical worker, who works under the direction of the writer and has been specially

trained to give and score the tests. As many as 160 applicants have been tested in one day. Applicants are tested singly, or in groups up to 10.

The tests are also used in cases of personnel adjustment, transfers, selection of candidates for training classes, etc.

A careful follow up by rating and production record is being carried on in order to check the test records with actual work. These are not yet in such a form that findings can be presented statistically, but on the whole, from a rough survey of the evidence at hand and from isolated cases which have come in for special attention the findings reported above tend to be borne out.

Summary. In a large department store methods of obtaining reliable criteria for work ability of salesclerks and clerical workers were carefully developed: Workers were grouped into three classifications, very good, poor and average, and tested with a series of mental tests; only those workers being tested who were *consistently* placed in one of these three classifications by all available criteria.

The good and poor salesclerks when tested gave negative coefficients of correlation, some markedly high. Similar groups of clerical workers likewise tested gave positive correlations, not so high.

When the two groups of workers, sales and clerical, were combined, the good workers of each type could be selected from the heterogeneous group by test marks, generally with accuracy as great or greater than from the homogeneous group. This was especially true if the poor workers of one group were omitted when selecting for the good of the other.

New tests were devised which were similar in form to those that gave the highest correlations. These tests differed in content by being less academic, more industrial. These and several tests in their original form were given to a group of 100 average salesclerks and 43 average clerical workers. The relationship between the two groups which had obtained with the original tests and the original groups tested held with the new tests and the new groups.

The regression equation was determined for the three tests which, taken singly, gave the best correlations—T 8, a Direction Test, T 12, a Judgment Test and T 13, a Comparison Test. The best weightings for these tests when combined were thus determined.

Biserial r for this group of tests and the group of 143 workers was .74. Twenty additional cases, clerical workers, raised the coefficient to .83.

These tests are now being used in connection with the employment of sales and clerical workers.

At a later date it is hoped to present data resulting from the follow up of the work of employees so tested.

In presenting this record of the work toward the development of tests for special ability, the writer is not unaware of the fact that much of the procedure and handling of the data varies from orthodox laboratory methods. Ideal laboratory conditions, are however rarely, if ever, possible to an industrial psychologist, who must make what shift he can. The exigencies of the occasion are therefore offered as an excuse for the departures from the orthodox. That in spite of necessarily improvised methods certain relationships were seen to hold consistently true, might in itself be offered as an "end justifying the means" excuse.

Also the subject of the negative correlations which were found, their why and wherefore, is a matter of some interest and importance which the present paper does not touch.

The findings themselves are presented, for comment and criticism, and with the thought that, in the light of the present scarcity of work on tests for special abilities, the greater scarcity of literature on the same subject, and a certain feeling of discouragement which has recently become associated with work in this field, the history of this work would possess some interest.

NOTE.—It should be noted in connection with the individuals tested in this experiment, that they possessed one attribute in addition to that of ability or disability at a special job, that attribute being *permanency* as workers. The workers were rated in June, sales records were begun from May, the testing lasted into January of the following year and ratings were unobtainable for workers who had not been employed for some length of time previous to June. Consequently only such individuals could have been included among those tested as were fairly permanent workers. Stability is as important a quality in a satisfactory employee as ability itself.

It is conceivable that any superior individual, making in the tests a score beyond the range of both clerical workers or salesclerks would do excellent work at either job. It is however (since promotion can of necessity come only to a limited number), highly inconceivable that such an individual would remain for more than the briefest period at such work as department store salesclerks or clerical workers are called upon to perform, day after day, and such individuals could hardly be counted on to make up the bulk of the 5,000 workers who carry on the business of such an organization.

The median intelligence of the workers in this store approximates, (from data not yet published), the degree of intelligence which was found characteristic in the Army Tests, and it is for such individuals that the standards here set forth hold true.

ESTIMATING INTELLIGENCE BY MEANS OF PRINTED PHOTOGRAPHS

By L. Dewey Anderson, Bureau of Personnel Research, Carnegie Institute of Technology

Purpose

The purpose of this investigation was (1) to determine the reliability of photographs for indicating the intelligence of strangers, and (2) to seek out differences in the ability of various persons to judge intelligence by this means. The results of the study might be applicable to problems of selection arising in an employment office.

Method

The photographs used were prints (from cuts) in a handbook published by a large department store company. They represented the faces of 69 employees of that company. These persons consisted of superintendents, buyers, managers and assistants.

In the fall of 1920 the same 69 persons were given the Bureau of Personnel Research adaptation of the army intelligence test, and the scores in this test were used as the intelligence criterion.

The photographs were all bound in the company annual, or handbook, but the printed matter on each page was concealed by a paper shield. The person judging the photographs, therefore, had to finger through the book, picking out the individuals, and designating them by the number which was given to each, to indicate their intelligence rank as described below.

Twelve judges passed upon the photographs. These judges were selected from a nearly homogeneous group of graduate students and instructors in psychology and there was very little chance of other influences entering into the judgments of intelligence given.

The instructions were as follows:

Instructions

"These photographs will introduce you to the 69 managers, buyers and assistants of the ——— Company. You are asked

to *estimate the intelligence* of these people by a careful study of their photographs.

"First, scan through the book and get an impression of the general character of this group. Then pick out the following:

(1) The seven most intelligent persons.

(2) The seven least intelligent persons.

(3) The fourteen persons who are superior, but not as good as the best seven.

(4) The fourteen persons who are inferior, but not as poor as the worst seven.

"The reason for the grouping by sevens is found in the fact that there are about seven individuals in each decile. The highest 10 per cent, the next 20 per cent, the lowest 10 per cent and the next lowest 20 per cent, are thus sought. This would leave an undifferentiated middle group of 40 per cent. The entire arrangement is thus intended simply to sift out the extremes."

TABLE I

SUCCESS OF VARIOUS JUDGES IN ESTIMATING INTELLIGENCE BY PHOTOGRAPHS

Judge	No. of Absolutely Right Judgments	No. of Absolutely Wrong Judgments	No. of Judgments on the Correct Side	No. of Judgments on the Incorrect Side
A	12	1	27	15
B	7	1	23	19
C	7	1	25	17
D	4	1	23	19
E	6	2	19	21
F	10	1	26	16
G	8	0	22	20
H	10	2	22	20
I	8	1	20	22
J	7	1	20	22
K	7	1	22	20
L	10	2	22	20
	96	14	271	231

In this table "absolutely right" means that the man's photograph was assigned to the *exact group* into which the intelligence test placed him. "Absolutely wrong" means that the very superior man was marked very inferior or the very inferior man was marked very superior. Placing the inferior man in the very superior group or the superior man in the very inferior group, etc., is indicated in the 4th column. "Judgments on the correct side" are those where superiority to the average has been rightly discriminated and where inferiority to the average has been rightly discriminated.

From the above table it will be noted that there are 96 absolutely right judgments and only 14 absolutely wrong judgments. This means that out of 110 judgments, 96 were absolutely right and 14 were absolutely wrong, or placed in the directly opposite group. In the last two columns of the table there are 271 judgments on the correct side and 231 on the wrong side. These data show that the judgments of intelligence based on photographs are, in this experiment, fairly accurate where only the extreme degrees of intelligence as measured by other criteria are involved. If all judgments are considered, i.e., if the middle group that is badly 'judged' is included, there is a very slight tendency to accurate placement.

A product moment correlation was made by assigning values to votes as follows: A judgment as very intelligent was given a rating of 2, a judgment in the least intelligent group a rating of —2, a judgment of superior, a rating of 1 and a judgment of inferior, a rating of —1. The ratings given to each photograph were summed and a product moment correlation of +.27 with intelligence test scores was secured. The probable error is .07. This correlation shows that there is a positive relationship between photograph ratings and intelligence test scores.

Another correlation was figured by the rank method between the judgments of 31 store managers on each other and the photograph ratings. The managers' judgments were made on the basis of general value to the firm or general ability as a store manager. A correlation coefficient of +.22 was found. This shows that there is some reliability in the correlation coefficients.

In all of these computations the "average" group is omitted.

Conclusions

I. *General Reliability of Photographs*

It seems very doubtful whether for practical purposes intelligence can be gauged by a study of a man's photograph. The correlation of +.27 between assigned ratings and intelligence as measured by the Bureau of Personnel Research Intelligence Test shows, however, that there is some tendency to discriminate intelligence above mere chance selection. From a study of Table I it was found that 16 men were placed at least once in both the most superior and the most inferior group, 49 people were placed in both the superior and inferior groups, 18 people were placed in both the most intelligent and the

inferior groups and 29 were placed in both the least intelligent and the superior groups. From these results, it appears that selecting employees by a study of photographs would be a very haphazard way of selection. When competent raters place men in both the highest and lowest classes 16 times out of a possible 69 chances, there is little possibility of upholding the practice of photograph rating.

There seemed to be considerable agreement among the raters in regard to some photographs. A point suggested by this agreement is that people in judging a person by his photograph usually pick out a particular feature such as the lustre of the eye or the width of the brow from which to draw their conclusion. In this study it was impossible to check this belief.

II. *Individual Differences*

There are considerable individual differences shown in ability to make ratings on intelligence from photographs. One rater made only four absolutely right judgments, while another made twelve. The average number of correct judgments is 8. From Table I, the number of absolutely wrong judgments is 14, thus showing, when compared to a total of 96 absolutely right judgments, that the raters were using some criteria of judging that served to pick out right men. The difference between poor and good raters is not as well marked in judgments that merely rightly place the photographs as in "absolute" judgments. The best record was made by "A," having 27 right judgments and 15 incorrect ones. The poorest records were made by two people, I and J, each having 20 correct and 22 incorrect judgments.

By dividing the number of correct judgments by the total number of judgments a percentage of accuracy was gained for each rater. These percentages follow:

A	64%	G	52%
B	55%	H	52%
C	59%	I	47%
D	54%	J	47%
E	47%	K	52%
F	62%	L	52%

TWO CASES SHOWING MARKED CHANGE IN I. Q.[1]

By W. T. Root, University of Pittsburgh

Every psychological examiner has, from time to time, certain cases in which testing and retesting even within a brief period of a few months yield decidedly different mental ages and consequently different I. Q.'s. Furthermore many of these cases cannot be traced to clerical errors, to moot questions in scoring, or to the personal equation, which may be a factor of discrepancy when different persons have administered the test at different times.

The cases here presented are typical of a not infrequent performance where school conditions permit individual attention and free grading. Both children came up for examination the second time due to the curiosity of certain teachers who had noted a marked improvement in the character of the work in each case, and also improvement in alertness and grasp. Perhaps the best way to get at these cases will be to give in detail the tests not passed the first time but passed the second time.

Case A

Age first examination, 9-0; mental age, 8-4; retardation, 8 months; I. Q., 92. Second examination, age 9-4; mental age, 9-9; acceleration, 5 months; I. Q., 104. Both times year VII was the base year. The first time the following tests were failed: year VIII, 1 (very poor), 2 (nearly performed), 5 (*b* is plus, *a*, *c*, and *d* failed); year IX, 1 (day of week only, given), 2 (nearly correct, 12 and 15 gram weights interchanged), 3 (very slow and inaccurate, not even good approximations, e. g., 25-4 equals 7), 4 and 6 (nearly passed). As IX 5 was the only one of this year passed, the testing would have closed with this year, but for the vocabulary score which carried the examination into the X year. In this year tests 4, 5, and A1 2 were failed. Tests 6 and A1.3 were not given. All of year XII failed.

The second time the examination was given the following tests were passed that were failed the first time: Year VIII, 1 (just scored), 2 (28″, 0 errors): year IX, 1, 2 (3 still failed 25-4), 6 (passed easily). The score for X remains unchanged

[1] From School of Childhood, University of Pittsburgh. Stanford Revision of Binet Tests used.

with the exception of A1 2 which is passed (barely) this time, but does not count in the score. In the XII year the interpretation of pictures is now sufficiently good to warrant scoring here.

In this case two factors offer plausible explanation for the advance in the second examination: (1) marked improvement in sight and the clearness of discrimination in reading; (2) rapid academic advance (noted by teachers) and most plausibly attributed to free and semi-individual method of teaching which permits a direct attack on the peculiar weaknesses of a pupil, in this case the three R's especially. A general clearing up in mechanics of reading, and a slight increase in verbal fluency and number combination would account largely for the differences in the two tests.

Case B

Age first examination, 6-8; mental age, 7-6; acceleration 10 months; I. Q., 112. Second examination, age 6-11; mental age, 9-8; acceleration, 33 months; I. Q., 140. In the first examination year V is the basic year; year VI, test 1 is failed; year VII, 6; year VIII, 1 and 6 (vocabulary 16); year IX, all but number 4 are failed, thus closing the examination.

In the second examination tests VI-1 VII-6, and VIII-1 and 6 are passed making VIII the basic year. The vocabulary is 22 words. There is a good performance of the Ball and Field against a complete failure the first time. In year IX tests 5 and 6 are passed with ease while 4 (repeat 4 digits backwards) is again barely passed. In year X, tests 2, 3, 5, and 6 are passed; in year XII, 3 and 7. Tests 2 and 6 (year X) are performed with ease and would probably have been passed the first examination.

In this case the increase in I. Q. can be accounted for plausibly by the factors suggested as causal for case A. Besides these factors which first suggest themselves, several other causes for fluctuation may be outlined tentatively:

1. A school permitting individual advance can make an attack on individual academic weakness, thus attention may be focused upon sentence structure, phonograms and addition combinations if they have proved to be weak. Such drill could conceivably have unintentional but nevertheless indirect effect upon year IX, 3, 5 and 6 (Making change, Using three words in a sentence, and Rhymes, respectively). These three tests correctly performed mean an addition of six years mental age. In the School of Childhood reading is delayed until the second grade, which might readily account for the failure in Rhymes

in the first grade, and passing of the test with ease after drill in *ay, ill,* and *ing* endings.

2. Every examiner has noticed that a child may just fail in several allied tests. Chance conditions may make these tests just barely plus on reexamination. Every teacher has also noted that after a rather discouraging plateau, a child may suddenly get a "hunch" in reading or number work. At times understanding seems to act on the *all or none* principle. The emotional attributes of failure and success could also play a part in test performance. (Case A and B, it will be recalled, came up for reexamination because their teachers had noted a marked improvement in their grasp of subject matter.)

3. Superior children are often those who have been brought under adult influence and culture to a marked degree. They possess a peculiar "adultish" ability to answer in a mature way verbal and moral questions, which gives a wide scatter range in testing, making it unsafe to discontinue the test per instructions. For example, all of year IX may be failed while test 2 and 5 year X and test 5 and 7 year XII may be passed. This is precisely what subject B did.

4. Improved physical condition may make an appreciable change in a short time.

Such cases as here described seem to indicate the need of more careful qualitative examination of test responses in terms of individual differences. Also there should be frequent reexamination where the nature of the school freely admits of more than average advance. And finally there should be frequent reexamination of children who have made marked physical improvement from any cause.

A PRELIMINARY STUDY OF THE CORRELATIONS BETWEEN ESTIMATES OF VOLITIONAL TRAITS AND THE RESULTS FROM THE DOWNEY "WILL-PROFILE"

By G. M. Ruch, Fellow in Education, Stanford University

The results summarized in the tables which follow were gathered in the attempt to compare the estimates of associates with the actual test scores earned in the Downey "Individual Will-Temperament Test" for the twelve traits comprising the scale.

Estimates were secured from two groups of associates, viz., university instructors and students in the same classes, for more than twenty advanced or graduate students. Except in two cases those rated had been known to the judges for at least six months and in most cases the judge and the student rated were fairly intimately acquainted socially as well as in academic associations. The numbers of judges varied from two to five for the faculty group and from four to eleven for the student group. No cases where the numbers of estimates secured were less than two or four, respectively, are included below; thus reducing the numbers concerned to fifteen except where a larger number is definitely indicated. The instructions given to the judges as to the nature of each volitional trait to be rated were taken verbatim for ten of the twelve from Dr. Downey's instructions in her paper, "Some Volitional Patterns Revealed by the Will-Profile" in the *Journal of Experimental Psychology*, Vol. 3, 1920, pages 281-301. The remaining two traits which have been added to a later revision of the scale, viz., "Finality of Judgment" and "Volitional Perseveration" were described by the writer in the instructions. It was suggested that the judges rate the entire group upon trait one before proceeding to the next trait in the hope that greater comparative discrimination would be exercised by those making the estimates.

The correlations were made by the Order-of-Rank method of Spearman corrected to approximate the Pearson Product-Moment formula.

The correlation between the total scores (sum of scores in twelve traits) and scores in Army Alpha for fifteen subjects was found to be —0.09. For the pooled estimates of faculty judges and Alpha, r was —0.12.

TABLE I

SHOWING THE CORRELATIONS BETWEEN THE SCORES FOR THE TWELVE TRAITS AND THE POOLED ESTIMATES OF THE JUDGES FOR THE SAME TRAITS, SUBJECT BY SUBJECT

Subject	Pooled Estimates of Faculty Judges with Test Scores		Pooled Estimates of Students with Test Scores		Pooled Estimates of Faculty with Pooled Estimates of Students
P.....	.26	(3)	.48	(11)	.88
Wy.....	—.06	(3)	.54	(10)	.16
Fe.....	.00	(5)	.31	(10)	—.65
Dar.....	—.15	(3)	—.13	(8)	.87
C.....	.02	(4)	—.13	(10)	.59
Wh.....	.13	(3)	.23	(10)	.53
R.....	—.18	(2)	—.64	(9)	.84
Dav.....	—.12	(2)	.57	(9)	—.04
B.....	.29	(4)	.10	(10)	.76
We.....	—.13	(3)	—.38	(10)	.76
Fa.....	—.11	(4)	.14	(7)	.83
L.....	.17	(4)	—.40	(5)	.76
H.....	—.13	(4)	—.05	(4)	.43
T.....	.03	(2)	.33	(10)	—.30
Deb.....	.10	(3)	.22	(7)	.70
Average.....	.01		.08		.47

( ) Number of judges.

TABLE II

CORRELATIONS OF THE ESTIMATES OF FACULTY AND STUDENTS AND THE TEST SCORES AND BETWEEN THE ESTIMATES OF THE TWO SETS OF JUDGES, TRAIT BY TRAIT

Test Number	Trait	Test with Fac. Est.	Test with Stud. Est.	Fac. with Stud. Est.	
1.	"Speed of Movement".....	.02	.43	.72	
2.	"Freedom from Load".....	—.09	.27	.28	
3.	"Flexibility".....	.45	.17	.22	(S)
4.	"Speed of Decision".....	—.02	.53	.67	
5.	"Motor Impulsion".....	.51	.50	.73	(S)
6.	"Assurance".....	.23	.37	.86	(S)
7.	"Resistance".....	.35	.28	.67	(S)
8.	"Finality of Judgment".....	—.11	.07	.45	
9.	"Motor Inhibition".....	—.33	—.23	.66	
10.	"Care for Detail".....	.15	—.07	.77	(S)
11.	"Coordination of Impulses".....	—.09	.05	.64	(S)
12.	"Volitional Perseveration".....	—.26	.05	.72	
Average.....		.07	.20	.62	

(S) Scoring in part subjective. All others purely objective.

TABLE III

AVERAGE INTER-CORRELATIONS OF THE THREE TYPES OR GROUPS OF TESTS WITHIN THE DOWNEY SCALE, VIZ.: TRAITS 1–4 OR "HAIR-TRIGGER" PATTERN; TRAITS 5–8 OR "WILFUL AGGRESSIVE" PATTERN; AND TRAITS 9–12 OR "SLOW, ACCURATE, TENACIOUS" PATTERN

	Pattern	Average Inter-Correlation of the Four Tests Making up the Pattern
I.	"Hair-trigger" (1–4)	—.005
II.	"Wilful Aggressive" (5–8)	.04
III.	"Slow, Accurate, Tenacious" (9–12)	.06

TABLE IV

THE RELIABILITIES OF ESTIMATES AND TESTS AS OBTAINED BY THE METHOD OF CORRELATING THE SCORES OR ESTIMATES FOR THE ODD-NUMBERED TESTS WITH THOSE FOR THE EVEN-NUMBERED TESTS, THE SCORES BEING THE TRAIT-SCORES

	r	N
Downey Scale	—0.15	15
	—0.21	22
Faculty Estimates	0.86	15
Students' Estimates	0.58	15

From table I it seems evident that the estimates of the judges show no relation to the actual test scores but that they do show a fair amount of agreement between the two sets of judges (average, .47). The students seem at first sight to be somewhat the better judges, but the apparent difference may be due to the fact that the number of students furnishing ratings was considerably greater than the number of instructors.

Table II shows a similar result. Considerable agreement between judges but little or none between estimates and tests is shown. It might be stated that students and instructors have a common concept of the significance of the traits (or rather of certain of the traits) but that these concepts do not correspond with the psychological processes concerned in the tests of the Downey scale. It was feared by the writer that his comparative inexperience in giving the tests, and particularly in scoring, might be responsible for the lack of higher agreement between estimates and tests. However, on the whole, those tests where the scoring is purely objective give poorer results than those where the personal equation does enter to some extent.

The treatment of data as given in table III calls for definite reference to an issue involved which may be entirely unfair to the Downey tests. The assumption has been made in

resorting to the method of average inter-correlations within the three groups of four tests each which have been designated as "patterns" that there should be a fairly high community of psychological processes within such patterns and that there would be a corresponding absence of close relation between tests of different patterns. This may not be a true conclusion, although Dr. Downey has stated that the tests are so grouped, tentatively, in order to bring together tests of the same general type.

Similarly in Table IV the same assumption has been made. The correlation of the odd-against the even-numbered tests really breaks the entire scale into half-scales in which each pattern is represented by two tests rather than four tests. This again may be unfair. However, the nature of two or more of the tests of separate traits makes it impossible to obtain measures of reliability by repeating the tests on another day. The increase in the number of subjects from fifteen to twenty-two lowers the coefficient somewhat rather than raising it.

The extremely small number of subjects used makes it imperative that all criticism of the Downey Will-Profiles be foregone even if the writer were so disposed. These results may possibly have some value and interest to others and it is hoped that they will serve to aid the efforts in a pioneer and inestimably valuable field of research. It is hoped to extend these preliminary findings as soon as possible to a point where greater reliability can be assured for the conclusions which may be drawn.

April 1, 1921.

PRELIMINARY REPORT ON A GIFTED JUVENILE AUTHOR*

.By Lewis M. Terman and Jessie C. Fenton

Data on gifted children have been accumulating at Stanford University since 1911. In 1915 one of us published a brief summary of facts relating to 54 children testing above 120 IQ, and certain additional material relating to 31 who had an IQ of 125 or above. (Terman: "The Mental Hygiene of Exceptional Children," *Pedagogical Seminary,* 1915, 22, 529-537.) In 1916, with the assistance of Margaret Hopwood Hubbard, more systematic work was undertaken in the locating, testing, rating of gifted children. In *The Intelligence of School Children* (Houghton Mifflin Co., 1920) fairly extensive data were reported on 59 cases, most of whom tested above 140 IQ. From 1917 to 1919 little was accomplished, owing to the interruption caused by the war, but during the last two years the work has occupied somewhat more than half the time of a research assistant. Our files now contain test records and other data on nearly 200 gifted children, most of them above 140 IQ. These are being followed up from year to year and as many as possible are given re-tests. At present the entire plan of work is being revised preparatory to gathering extensive psychological, educational, social and biological data on 500 to 1,000 of the most gifted children in California. An important feature of the plan is to follow up as many of these cases as possible for a period of at least ten or twenty years. One of the most interesting of the children thus far studied is the gifted juvenile author described in the following pages.

Betty Ford was born in San Francisco, January 21, 1912. Her four grandparents were of Swedish, German-French, English, and Scotch descent, respectively. There is nothing especially remarkable in her family tree. The mother is a woman of more than average intelligence and of considerable musical ability. The father is a physician and the author of the "Ford Stitch," favorably mentioned in standard texts on surgery. Betty is an only child.

*Read before the American Psychological Association, December, 1920.

In Betty's baby-book, kept by her mother, we find the following entries:

Weight at birth, 11 pounds, 15 ounces.
Weight at 20 months, 31 pounds.
At 7 months began to walk about in her "walker."
At 19 months said everything clearly and knew the alphabet.
At 20 months put together block puzzle pictures.
At 22 months enumerated birds up to 12 in a picture.

In view of her intellectual precocity it is interesting to note that Betty's earliest memories, if we can accept her own testimony, date back to the time when she was still fed by bottle, at the age of about 12 months.

Betty's parents say that they first learned of her ability to read when they discovered her, at the age of 4½ years, reading Heidi, a book of about fourth grade difficulty. At five she read such books as Silas Marner and Charles Lamb's stories. By her eighth birthday she had read approximately 700 books, many of them twice. For the last two years she has read on an average at least eight hours a day. She reads rapidly and with complete absorption. Instead of finishing one book before proceeding to another, she prefers to keep three or four going together. At the age of eight her favorite authors included Shakespeare, Burns, Longfellow, Wordsworth, Scott, Poe, Victor Hugo, Thackeray, Dickens, Cooper, Mark Twain, Barrie, Conan Doyle, Kipling, R. L. Stevenson and George Eliot.

At the age of 30 months Betty began making up little stories about rabbits, frogs, fishes and squirrels. At 33 months she composed little jingles in rhyme. Stories and jingles continued to appear from time to time, but before the sixth birthday only a few were recorded. When she was six years old she was given a typewriter, and with its assistance and stimulus she was soon well launched on her literary career. By the age of eight she had written more than 100 poems and more than 75 stories, together enough to fill a volume of 200 or 250 pages.

We first saw Betty in November, 1919, when she was 7 years, 10½ months old. Since that time we have spent approximately thirty hours with her, giving her tests, showing her about the University or visiting with her in her home. Six weeks before her eighth birthday she tested 14-10 by the Stanford Revision; IQ 188. Her vocabulary at that time was approximately 13,000 words, which is not far from the median for Stanford University freshmen. Her score on the Army

Beta test was 71, which is about the norm for 14 years. Thus her IQ measured by a non-verbal scale is 175. By the National Intelligence Tests she graded above 15 years (score, Form A 152.8, Form B 123), and by the Terman Group Test considerably higher (score, 151, Median for Grade 12). On the Trabue Completion tests, B and C, her score (17.5) equaled the median for third or fourth year high school pupils. She took the memory span test for digits (oral presentation) with a class of 21 graduate students, beating ten and tying four. Her extreme memory span is 9 digits, direct order, and 7 digits, reversed order. On the Kelley construction test, which was given her by Dr. Kelley, her performance resembled that of a normal child of nine or ten years. On the various types of tests found in the Stanford Revision, the National, the Terman Group and Beta, her performance runs fairly even, though not perfectly so. She is at her best in vocabulary, information, language completion, and memory, and at her worst in tests involving arithmetical processes.

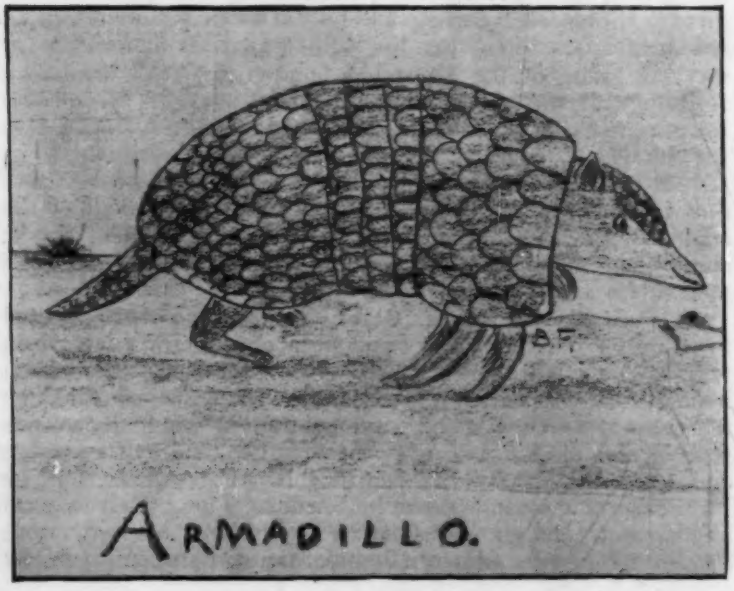
Betty has never attended school, but for nearly two years she has had about two hours a week of private instruction in dancing, French and German. During the last year her mother has given her, intermittently, a small amount of instruction in arithmetic, usually not more than an hour a week. On the Woody tests for addition, subtraction, multiplication and division her scores correspond to the medians for grade 5. On the Kansas Silent Reading test she scores at median for the 11th grade, on the Thorndike handwriting scale at median for 8th grade, on the Starch Grammatical Scale (A) at median for 8th grade, on the Starch Punctuation Scale (A) at median for 6th grade, on the Teachers College Spelling Scale (List 16) at median for 8th grade, on the Grier tests of information in zoology and physiology considerably better than the standards given by Grier for second term high school pupils, and on the Abbott-Trabue test of poetic appreciation at the median for second year college students.

As a test of Betty's poetic ability, five of her poems, selected by a dozen judges as her best, were chosen for comparison with twelve other poems, including the following: three from a published volume of poems written by Stanford students, three by Longfellow (one of them written when the author was 9 years old and one when he was 13), one by Tennyson, one by Blake, one by Blanden, one by Wordsworth, one by Herrick, and a juvenile poem by Shelley. These twelve poems and five of Betty's with names of authors omitted, were arranged for order of merit by 46 Stanford students (chiefly

seniors) enrolled in advance classes in versification, Browning, and composition. In the combined rank order, based on the ratings of the 46 judges, Betty's five poems held ranks 8, 9, 10, 11 and 14. For each of her poems the mean deviation of the 46 rankings from the average rank was between 2.5 and 4.1.

The rank orders based on the combined ratings were as follows:

1. Monterey: Stanford poem by Glenn Arthur Hughes
2. Song: Stanford poem by Dorothy Gunnell Jenkins
3. Shaneen: Stanford poem by James Leo Duff
4. Snowflakes: Longfellow
5. The Eagle: Tennyson
6. The Battle of Lovell Pond: Longfellow (Written at age 13)
7. The Song the Grass Sings: Blanden
8. My Prayer: Betty Ford
9. After the Rain: Betty Ford
10. The Clouds: Betty Ford
11. Joy: Betty Ford
12. March: Wordsworth
13. The Wildflower's Song: Blake
14. The Rainbow: Betty Ford



15. To Violets: Herrick
16. Mr. Finney's Turnip: Longfellow (Written at age 9)
17. The Cat in Distress: Shelley (Written at age 8)

The selection of poems for rating demands a word of explanation. In general, poems were selected which were not too unlike Betty's in content, which were presumably not well known by the judges, and which were relatively brief. Three were selected because they were juveniles. The five of Betty's were selected by about a dozen graduate students as being among her best. Three of the five were judged by Betty as among her best.

It will be noted that the first three places are won by Stanford poems. This may be entirely due to the fact that they were probably known to a good many of the judges, though they are not without considerable merit. Two of the juvenile poems by well known authors are rated at the bottom of the list, but one of the juveniles by Longfellow is given fourth place. We reproduce some examples of the seventeen poems, those of Betty in their original spelling and punctuation.

SONG

(Stanford poem. Rank 2)

There is life in the green of the hills,
  There is joy in the mist in the hollow,
Where the song of the meadow lark thrills,—
  Up, my heart, follow.

There is joy in the song of the sea,
  There is life in the dip of the swallow,
And the voice of the world calls to me,—
  Up, my heart, follow.

THE EAGLE

(Tennyson. Rank 5)

He clasps the crag with hooked hands;
Close to the sun in lonely lands,
Ringed with the azure world, he stands.

The wrinkled sea beneath him crawls;
He watches from his mountain walls,
And like a thunderbolt he falls.

MY PRAYER

(Betty Ford. Rank 8. Written at age 7 years, 6 months)

Oh, Master of fire, Oh Lord of Air,
Oh God of waters, Hear my prayer.
Oh, Lord of ground and of stirring trees
O God of Man and of pleasant breeze
Dear Father, Let me happy be
  As happy as a growing Tree.

AFTER THE RAIN

(Betty Ford. Rank 9. Written at age 8 years, 10 months)

Birds carol glad and free,
All about joy you see,
After the rain!

Each little throat swells glad,
None can be ever sad,
After the rain!

Moss is a velvet green,
Grasses of silken sheen,
After the rain!

Blue is the sky above,
Everything smiles with love,
After the rain!

Birds serenade the sun,
Each little songster one,
After the rain!

Flowers smile back at him,
No one can e'er be grim,
After the rain!

And so let Nature prove,
In all is God's great love,
After the rain!

THE CLOUDS

(Betty Ford, Rank 10. Written at age 7 years, 3 months)

The Clouds are soft white feather beds
On which the fairys lie
the fairys all look down on us
as they come floating by

but when the clouds grow angry
they throw the Fairys down
The fairys are so happy
to visit at our town.

Ques.
What are the fairys?
Ans.
Raindrops!

THE RAINBOW

(Betty Ford. Rank 14. Written at age 7 years, 3 months)

The rainbow of Diana
  Is shineing in the sky
The sunbow of Apollo
  has not just yet come by.
  Oh shine again
thou lovely rainbow
  the dark'ning
Twilights coming now
thy brother's waiting
  O'er the sea
He hath not yet come
  home to thee.

A CAT IN DISTRESS

(Shelley. Rank 17. Written at age 8)

A cat in distress,
  Nothing more nor less:
Good folks, I must faithfully tell ye,
  As I am a sinner,
  It waits for some dinner
To stuff out its own little belly.

  You would not easily guess
  All the modes of distress
Which torture the tenants of earth;
  And the various evils,
  Which like so many devils
Attend the poor souls from their birth.

  Some a living require,
  And others desire
An old fellow out of the way;
  And which is the best
  I leave to be guessed,
For I cannot pretend to say.

  One wants society,
  Another, variety.
Others a tranquil life:
  Some want food,
  Others, as good,
Only want a wife.

  But this poor little cat
  Only wanted a rat,
To stuff out its own little maw;
  And it were as good,
  Some people had such food,
To make them hold their jaw!

Betty can sometimes be persuaded to improvise a story or poem on a subject selected by another person. In such cases she dictates it, often more rapidly than one can take it down in longhand. One improvised poem of 16 lines she dictated to one of us in 2 minutes and 50 seconds. Twenty-one university students in an advanced class in versification were allowed fifteen minutes in which to write a poem on the same subject. This test was given by the regular instructor as a class exercise. The poems were then typed and given to 35 judges to be ranked in order of merit. In the combined ranking Betty's poem held sixth place. Besides Betty's, we reproduce the best and the poorest written by the university students.

A TALE OF OTHER BLACKBIRDS THAN THE ONES IN THE KING'S PIE

(Betty. Written at age 8)

Four and twenty blackbirds
Seated on a roof
One did peck his neighbor
His neighbor grunted, "OOf".
The neighbor also, he got mad
And pecked another bird,
His neighbor also wasn't glad
And said a naughty word.
This neighbor pecked another yet,
The neighbor vowed he'd not forget;
took vengeance on another fowl
And raised a long and doleful howl.
Then that one turned without a word
And pecked the very lastest bird,—
So all the birds got in a quarrel
Because of one ungodly fowl.

Moral: If you do good, the good will be mirrored in those around you. Also, if you be bad, the bad will be imitated by those around you.

BLACKBIRDS

(Rated poorest among the verses by university students)

I sat on a window sill,
The sun was shining bright;
A row of blackbirds chattering, fill
My mind with thots all right.
But to tell what I have in mind,
Is harder by far, than to guess
What the twitter of those birds mean,
As they spatter their words about.

I gaze on their pointed faces,
And see by twirls and grimaces,
The creatures are poking at us,
We who think we are smart,
At our expense and display,
Fun of the grandest sort.
One calls us worms, creatures that live in the clay,
Another like grass dried up like hay.

BLACKBIRDS

(Rated best among the verses by university students)

Out there on the roof in the sunshine
  Twelve blackbirds are stopping to rest.
Their merry chirping and chatter
  Perchance, speak of spring and a nest.

We sit at our desks near the window,
  And painfully, patiently try
To keep our eyes fixed on a lesson—
  Our minds on some old "Poly Sci."

Oh, why cannot we, like the blackbirds,
  Sit out in the bright morning sun
And muse on the coming of summer
  When all of our lessons are done?

But we cannot fly like the blackbirds
  Far into the deep, azure sky—
Instead,—we may soar after knowledge
  And mount after learning so high.

The majority of Betty's poems have been composed in less than ten minutes. The sight of a rainbow or sunset, or robin, perchance some experience while at play, or a suggestion from her reading, brings a flash of inspiration which sends her post haste to her typewriter. Recently, however, she has come to prefer dictation, because it leaves her freer to think. Hardly a word is changed during dictation and few afterward. Few false starts are made. At times she dictates more rapidly than anyone can write in longhand. While dictating her voice is low and her words run into each other so that they are hard to catch. Meanwhile she walks, skips or dances about over the room. Her stories are composed much less rapidly than her poems, perhaps at an average rate of 300 or 400 words in a half hour.

Betty's prose compositions have not been rated, but on the whole they seem to be about as remarkable as her poems. Both stories and poems show a prolific imagination, a lively sense of humor, and keen powers of observation. Her diction is usually good, but her sentence structure often shows carelessness and in poetry her meter sometimes limps. We repro-

duce a brief prose poem, a few lines of description from a story entitled "The Honeymoon," a fairy story, and one from a series of nature stories entitled "Stories the Alphabet Told" (one for each letter of the alphabet, written as a Christmas present to us).

FAIRY DEFANITION.
(Written at age 7 years, 11 months)

Fairies are the fancies of an imaginative brain,
Which wearying of earthly realities aspires to
create beings liveing only in thought
endowing the spirits thus created
with all genius for giving
HAPPINESS.

P. S. I wonder why grown-ups scoff at faries
And believe that angels have wings.
I don't think one is any harder to
Believe than the other,
Do you?

PARAGRAPH FROM "THE HONEYMOON."
(Written at age 7 years, 10 months)

"They stopped a moment to watch the sunshine glistening on the stream and to watch the suckers, trout and minnows darting through the water. Once they caught sight of an eel. The forest was so interesting that day. At last after gazing a long time at the silvery fish they again entered into the wilderness. A little bunny scurried across the path, a robin chirruped, a little squirrel came so close they were almost afraid of him, a cow mooed afar off, a pair of grouse scurried across the road even as the rabbit had done. A keen looking fox slunk along to the east. Once they heard a bear growl. Once a gentle eyed fawn rushed close past them. A pair of quail with frightened calls rose from some brush and flew over their heads to the other side. Once they came across the nest of some baby goldfinches."

WHY THE DAISY HAS A GOLDEN HEART AND ROSY PETALS.
(Written at 7 years, 2 months)

ONCE the Daisy had no heart. Not even a sign of one. One day as she was loudly bewailing her fate like this—Ah me, Would that I had even a tiny stone to take the place of a heart. Would that—At that moment the Sun came by.

What are you alasing about little one? he said with a jolly laugh. Alas the daisy replied, I am weeping because I have no heart. Ho, Ho, laughed the Sun. "Is that all"? THAT is easily remedied. You know the sun is all made of gold, so he pulled out some of his hair and fashioned it into the loveliest little golden heart you ever saw and laid it in the breast of the daisy. There little one, said the jolly sun. There take that and whenever you'r glad that you have a heart think of your old friend, the Sun, and he sailed away. (But the dear little daisy was modest and so timid and shy that what do you suppose she did,—She folded her petals tight around her golden heart. But in the morning she grew a little bolder and opened her pretty petals and all the flowers said, how wonderful the daisy has a golden heart.

End of part I

PART II WHY THE DAISY HAS ROSY PETALS.

Little Daisy stood in all her beauty and the sun came by, there was a little butterfly on a sunbeam—snow-white with spots of gold like the daisy herself—So she called out—

"Little butterfly dearest come down and talk with me. But the butterfly did not hear. She called a second time—Then the butterfly heard and foulding his wings dropped right down on the daisy's petals. Oh you beautiful thing he said, you beautiful thing.

And the Daisy blushed for joy—and ever since then the daisy has her rosy tinge.

Moral. the things we hope for may come to pass.

STORK

(Written at age 8 years, 10 months)

S stands for Stork who brings babies 'tis said,
But before any see him, he's back safe in bed.

Sallie Stork and Sammy Stork were heart-broken. They had been married some weeks ago, they had built a nest, but Sallie had laid no eggs as yet. "Oh dear," they bewailed in concert, "What will become of us if we have no children? Oh dear, oh dear, we must have some children, we must. But at last one day, Oh joy, could it possibly be? There in the nest lay a big, white egg. What should she do! Oh they were in paroxysms of delight.

They waited a long while and happy day, who came out of the egg but little Susan. Oh happiness! "Mamma," she inquired, in a sweet stork lisp, "Where am I? What's happened to me? The last thing I remember is good St. Nicholas taking me to one of the angels who put me in a tight white crib. What happened then, mamma, wont you tell me?"

"Well," her mother answered, "I don't exactly know, but it appears to me that when the angels put you into the egg, that's what it seems to me the crib was,—they brought you down and put you in our nest."

"Well, that's nice," said Susan, I'm sure I couldn't have found a better mother in heaven. You know there aren't generally any storks up there."

"Aren't there?" her mother answered, "Well I'm glad of that for perhaps if there had been many, you wouldn't have come down to us, you darling." "Indeed," chipped in Father Sammy, "I don't believe they could have brought us a finer young Stork.

So the days passed happily along and Susan grew tall and big and soon learned to fly.

At length one day she suggested to Sammy and Sally, "Why cant we take a trip up to Heaven now, I'm sure my wings are strong enough, why not?"

"Very well," Father Sammy answered, and Mother Sally agreed, "If you want to we will do it. When shall we start?" He enquired of her.

"Why not start to-day?" suggested their neighbor, Susie, brightly, who had come in for a little chat.

"Why not?" answered Susan, who was no longer little, "So here we go," and they soared high up in the sky, Susie included.

Finally they got there. All at once Susan gave a cry of joy—"St. Nicholas," she cried, and flew rapidly to-ward him. He was a jolly roly-poly fat man. "Why if this isn't Susan!" he cried gladly.

"My but you've been away a long time. But we are in great trouble," he went on, "I really dont know what to do. The angels are all busy taking down babies or watching over others and there are four human babies who have just arrived and have to be taken to their mothers. What shall we do? You know we're all occupied. The worst is that one cant take them alone, or I might do it, but they live in different parts of the world, one in Chicago, one in Hong-Kong, one in Amsterdam and one in Petrograd; and they all have to be delivered to-day C. O. D. The charge is 10 Lbs. of happiness C. O. D. Except the mother in China, she only pays five because they dont love girl babies in China as much as the boys. But as its the only Chinese vociferous here, it will have to go, I suppose."

"I'll take the Chicago one," offered Susan.

"That's nice," cried old Santa. "Very obliging of you I'm sure. The address is, 1589 Lover Lane."

"Dont count me out of the game," intercepted Susan. "I'll pack the squaller of China."

"Well that address," said old Santa, "is XYZ Ching Chong Joy Path."

"I'm sure I'll help too," remarked Mother Sally. "Give me the little Dutchman. I've heard that little Dutch babies aren't as noisy as the others."

"Oh yes they are," St. Nicholas laughingly assured her, "But if you wish you can take charge of that fellow," and she flew off to get the baby.

"Well I guess I'll take a chance if the others do," Father Sam decided "And as the baby Russian is the only one left, I'll take that."

"All right, that's on Lovesky Lanesky. Deliver it to the Countess Cuddlesky," and stooping with difficulty he pasted on the youngster's stomach, a label bearing, "Glass,—With care," and Sammy flew off.

"There" said old Santa drawing a big sigh of relief, "I do wonder when they'll come back."

In a little while Susan came back. "Gracious goodness, that woman gave me 20 Lbs and didn't ask for the change," she remarked as she settled down beside him.

Old Santa laughed with satisfaction. "Well she was certainly generous she must be a millionaire at the joy bank."

In another minute back came Susie. "You said that woman should only be charged 5L. but she handed over 25L. Let that 'not loving' stuff be forgotten, why I bet if I had waited long enough she'd have given me five million.

And then came Sally. "Ten pounds did you say? She handed over ten hundred it seemed to me. Gracious they must have a million millions of happiness by the way they spend it.

At last came Father. "Aren't the Russians generous? You said ten pounds, didn't you? They gave ten thousand and offered to give more if I'd wait a little.

"I've enjoyed the trip so much," Susan timidly hinted,

"That I think we might" (from Susie)

"Be the Saint's (this from Sallie)

"Messengers." (this from Father Stork.)

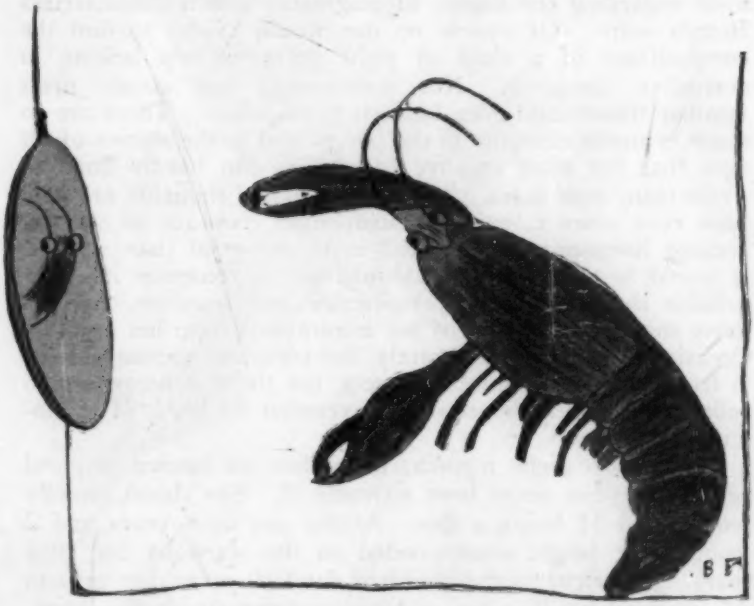
"That's right," and they looked at each other in satisfaction. "That's right" Old St. Nicholas approved, "We'll take you into our service," And ever since the storks have had no occasion to be idle!

Moral. The only blissful joy, you know,
Is, in making some other one's happiness grow.

Stork. Family—Ciconiidae, are large and white except the extremities of the wing tips, which are black. Their food consists of frogs, birds, reptils and insects. Storks build their nests on chimneys, trees or high rocks. The eggs are white. Its common name in Dutch means, 'Bringer of Good.'

Many of her poems and stories she has illustrated with colored crayon or pencil drawings. Some of these give evidence of considerable talent. Her pictures are never copied directly. If she needs a model (as for some animal picture) she finds as many different pictures as she can and then makes her own, using the best features from each model, together with such changes and additions as suit her purpose. She draws very carefully, making many corrections; then, when she is at last satisfied, she traces her final outline with carbon paper in order to secure a neat copy, and colors it with crayon.

Very frequently she appends to story or poem a "moral," well pointed and aptly expressed. One not knowing her might attribute this moralizing tendency to priggishness, but



LOBSTER.

it is really due to her playful enjoyment of the intellectual gymnastics involved in generalizing a situation or incident.

Although Betty's compositions show a wide range of interests, there is a marked preference for themes relating to nature. Classification of 86 poems gave the following result: nature poems (descriptive and appreciative), 44; reflective poems, 11; prose poems, 8; homely poems, or poems of everyday affairs, 7; love poems, 6; occasional poems (Christmas, election-day, etc.), 4; plays and extravaganzas (one a play of 1,500 words), 3; unclassified, 1. Of 65 prose compositions, 31 are nature stories of the *pourquoi* type ("Why the bunny has a white cotton tail," "Why the daisy has a golden heart and rosy petals," "Why the owl loves the moonlight," "Why the thistle has thorns," etc.), 25 are animal stories (in a collection entitled "Stories the Alphabet Told"), 6 are fairy tales, 1 a narrative story of 1,500 words (entitled "The Honeymoon") and 1 a detective story of 2,500 words.

It is not possible at present to make any quantitative statement regarding the degree of originality which characterizes Betty's work. Of course no one would expect to find the compositions of a child of eight years entirely lacking in derivative character. Not infrequently one comes upon familiar themes and even familiar phraseology. There are so many common elements in the poetry and in the stories of all ages that the most creative of writers can hardly hope to strike many new notes. Absolutely original thoughts are perhaps even more rare in literature than they are in science, because literature is older and more universal than science. It would be psychologically absurd not to recognize it as inevitable that such an impressionable and sensitive child as Betty should draw many of her inspirations from her reading. Occasionally, though very rarely, her voracious memory seems to have swallowed entire sentences, but there is no reason to believe that her borrowings have exceeded the limits of unconscious absorption.

Except for slight myopia, Betty has no known physical defects, and has never been seriously ill. She sleeps soundly from 10 to 11 hours a day. At the age of 8 years and 2 months her height corresponded to the standard for 10½ years, her weight to the standard for 9½ years, her grip to the standard for 10 years, and her lung capacity to the standard for 8½ years. It is perhaps because of her unusual activity that she prefers to play with boys rather than with girls. In doll play and other ordinary occupations of girlhood

she has shown only moderate interest. Her ambition is to be a physician and a writer.

Ratings of Betty on 46 physical, intellectual, emotional, volitional, social and moral traits were secured from the mother and from a school teacher who had known the child intimately since babyhood. The ratings were made on a scale of 1 to 5, 1 being defined as "very superior," 2 as "superior," 3 as "average," 4 as "inferior," and 5 as "very inferior." The average of the mother's ratings was 1.43; of the teacher's ratings, 1.47. The mother's ratings included only two as low as "average" (those on intellectual modesty and communicativeness). The only rating by the teacher as low as "average" was that for beauty.

Our own rating on the traits which we have had a chance to observe would not differ greatly from those of the mother

or teacher, although they might be somewhat lower on intellectual modesty and social adaptability. In neither of these traits, however, does there seem to be any marked inferiority, if one makes due allowance for the difficulties which are inevitable in such an extraordinary case of intellectual precocity. It is probably this, chiefly, which explains the fact that she prefers the companionship of children who are either several years older or several years younger than herself. The mother's habit of restricting her associations with children of her own age has doubtless tended to confirm this natural but not altogether wholesome preference. There is no evidence that when she does play with other children there is any marked difficulty of social adjustment. Doubtless Betty knows that she has talents which other children do not have. She is too intelligent to be kept in ignorance of the fact. But if she is conscious of her superiority she gives little evidence of it. The mother has tried to imbue her with the idea that if she has been endowed with any unusual gift it is her duty to use it to make the world happier and better. "Mother," she said one day, "I am not proud because I can write verses; they just come themselves. But if I could only learn to control my temper, then I would be proud."

Betty's education has been managed by her mother, whose guiding principle has been the conviction that a child's confidence in its own ability to think and do should be sedulously cultivated. She believes that the abilities of children often atrophy for lack of encouragement or because of the destructive criticism of their elders. When Betty does a thing well she is commended. Spontaneity is cultivated. Her literary ability cannot be regarded as an artificial product of special training. It was her mother's ambition that she should become a musician. When it became evident that the child's talents lay in another direction she was encouraged to follow her own bent.

ON THE NEED FOR CAUTION IN ESTABLISHING RACE NORMS

By Ada Hart Arlitt, Bryn Mawr College*

In recent years there has been a tendency to establish race norms on the basis of mental tests given to representatives of various races who have emigrated to the United States. The results of tests made of groups of children of foreign born parents have been stated and comparisons made with the distribution of intelligence in groups of children of American born parents. So far as we have been able to ascertain except in the case of negro children[1] no attempt has been made to check up the influence of factors, other than that of race, which might quite as well account for the differences which it has been assumed were traceable to race alone. Particularly is this true in regard to the part played by social status. Much has been written about the effect of social status on the median intelligence in any group of American children, little or nothing has been said as to how much this may account for the low or high median score or median intelligence quotient in groups of children of other races. For instance Terman[2] speaking of the influence of race on Intelligence Quotient cites the median I.Q. for Italian children as being 84 or 16 points below that usually given for American children.

It was with a view to determining the relative influence of the two factors, race and social status, that this investigation was begun.

The tests used were the Stanford Revision of the Binet tests. These were given by the author and the four graduate students whose names appear as collaborators. All of the students were experienced testers. In order to insure uniformity of grading all tests were marked in consultation with the author.

The children 343 in all were taken from the primary grades

*(With the assistance of C. Baechle, I. Neterer, K. Smith, and C. Barnes.)

[1] Sunne Dagney. A Comparative Study of White and Negro Children. *Jour. of App. Psychol.*, 1917, p. 71.

[2] Terman. Intelligence of School Children, p. 56. (Houghton Mifflin Co., 1919.)

in a single school district. Of these 191 were children of native born white parents, 87 were Italians and 71 were negroes. All of the Italians spoke English without difficulty.

The native born white children were divided on the basis of social status into five groups with reference to the occupation of the father. The separation into groups by occupation followed Taussig's division into the five non-competing groups, i.e., (1) Professional classes; (2) Semi-professional and higher business; (3) Skilled; (4) Semi-skilled; (5) Unskilled. The home conditions followed closely the division by occupation. The last two groups, children of unskilled and semi-skilled workers, contained too few children to be treated separately. They were therefore combined. As a matter of convenience these children will be called throughout the "Inferior Group," the children whose parents fall into the skilled class will be called "Average," those of the semi-professional and higher, business class "Superior" and those of the professional class "Very Superior."

Of the Inferior group there were 43, of the Average group 76, of the Superior group 48 and of the Very Superior group 24.

TABLE I

Social Status	50–59.9	60–69.9	70–79.9	80–89.9	90–109.9	110–119.9	120–129.9	130–139.9	Above 140	Total
V. Sup.	0	0	1	0	6	4	6	3	4	24
Sup.	0	0	0	0	12	17	14	4	1	48
Av.	0	1	3	2	41	16	10	2	1	76
Inf.	1	3	5	9	23	1	0	1	0	43

The effect of Social Status on the distribution of Intelligence Quotients.

The outstanding characteristics of each group can best be seen from Table I. Over 40% of the Inferior social status group have I.Q.'s below 90 and only 4.6% have I.Q.'s above 110. There are only two children of very superior mentality in this group. In the Average group 50% are of average mentality and over 25% are above average i.e., have I.Q.'s above 110. The curve for the Superior group shows it to be as much better than the Average group as that was superior to the Inferior social status group. The median I.Q. is 118.7 or 10.2 points above that of the Average group. Seventy-five per cent of these children have I.Q.'s above 110 and none have I.Q.'s below 90. A curve for the Very Superior group

would have two modes, one at 90-110 the other at 120-129. Seventy per cent of these children have I.Q.'s above 110 and of these 76% have I.Q.'s above 120. The difference between the Superior and Very Superior social status groups lies not in the proportion of the children who reach 110 or above but in the fact that a larger proportion of those whose I.Q.'s pass 110 reach 120 or above.

It is interesting to note that very few children of superior mentality were found in the Inferior social status group and that only one child of inferior mentality was found in the Superior and Very Superior social status groups.

The median I.Q.'s for the four groups were respectively 92, 107, 118.7 and 125.9 or a difference of 33.9 points between children of Inferior and Very Superior social status of the same race and attending the same grades in the same school.

We are well aware, as has been stated, that previous writers[3] have emphasized the part played by social status. It is here cited chiefly because the variation due to social status is wider than that usually reported and serves to demonstrate all the more clearly the part that this factor may play in determining the distribution of intelligence in any non-selected group.

The effect of social status on the native born white group having been determined the total group was divided on the basis of race alone. It is a significant fact that there were too few children in the Italian and negro groups of even average social status to enable us to study the combined effect of race and social status in these groups.

The median I.Q. for the native white group was 106.5,[4] for the Italian group 85 and for the negro group[5] 83.4. The native born white group had a median I.Q 21.5 points above that of the Italian and 23.1 points above that of the negro group.

[3] Wientrob, J. and R. The Influence of Environment on Mental Ability, *Jour. of Educ. Psychol.*, 1912, pp. 577-586. Terman, L. M. The Stanford Revision of the Binet-Simon Scale, Chap. V. (Warwick and York, 1917.) Terman. Intelligence of School Children, pp. 188-189. (Houghton Mifflin Co., 1919.)

[4] The high median I.Q. is probably due to the fact that most of these children were between the ages of five and nine. The range of the I.Q.'s made by the middle fifty per cent of the children tested by Terman was from 97 to 111 in five and six year olds combined, and from 95 to 111 in seven and eight year olds combined. The average child might therefore be expected to make a higher I.Q. at these ages than at later ages as for instance thirteen to fourteen where the limits of the quotient including the middle fifty per cent are from 90 to 105. (Data cited from "Stanford Revision of the Binet Tests," page 40.)

[5] The negroes were of mixed blood. It was impossible to determine the proportion of white blood in the case of any of these children.

This disparity is great and if it had been due to race alone would have demonstrated beyond a doubt the superiority of the native white group, but 37% of the native white group came from families of Superior and Very Superior social status whereas 93% of the negro and 90% of the Italian group came from families of Inferior and Very Inferior social status.

The Italian and negro children were then compared with the group of native born white children of Inferior and Very Inferior social status. The median I.Q. for the native white group was then 92 or 8.6 points above that of the negro and 7 points above that of the Italian group.

This difference is not due to a larger proportion of children of superior or very superior mentality in the native

TABLE II

	50–59.9	60–69.9	70–79.9	80–89.9	90–109.9	110–119.9	120–129.9	130–139.9	Above 140	Total
Native born White.....	1	3	5	9	23	1	0	1	0	43
Italian.....	3	11	19	17	26	4	1	0	0	81
Negro.....	1	7	16	20	25	1	1	0	0	71

The effect of Race on the distribution of Intelligence Quotients.

white group. There are five such children in the Italian group, two in the negro and one in the native born white group. It is due to the proportion of children in the negro and Italian groups whose I.Q.'s are below 80. In the native born white group 20.9% have I.Q.'s below 80, in the negro group 33.8% have I.Q.'s below 80 and in the Italian group 41.2% fall below 80. As all of these children live in the same section of the same community and attend the same grades in the same schools this difference both as to the median I.Q. and as to proportion of dull children seems to be racial. But the difference in median I.Q. which is due to race alone is in this case at most only 8.6 points whereas the difference between children of the same race but of Inferior and Very Superior social status may amount to 33.9 points. It is apparent that such differences as we have between the negro and Italian children and between these

[6] These three groups were originally the same size namely, 29 in each. Other cases have been added to bring the groups up to their present size with a consequent change in the median I.Q. of only .5 of a point. The median I.Q., therefore, seems to represent a difference which is not due to the difference in the size of the groups.

and children of native born white parents are not nearly so striking as the difference between children of the same race but of different social status. Of the two factors social status seems to play the more important part. To such an extent is this true that it would seem to indicate that there is more likeness between children of the same social status but different race than between children of the same race but of different social status.

While the number of children tested is too small to permit us to make statements which may be universally applicable our results seem to warrant the following conclusions:

1. There is a difference in the distribution of intelligence in negro and Italian groups as compared with groups of children of native born white parents which seems to be due to race alone.

2. There is a marked difference in the distribution of intelligence in groups of children of the same race but different social status. This disparity seems to be greater than the disparity between children of the same social status but different race.

3. Race norms which do not take the social status factor into account are apt to be to that extent invalid.

THE PREDICTIVE VALUE OF SHORT INTELLIGENCE TESTS

By C. F. HANSEN and M. J. REAM, Carnegie Institute of Technology

Just how long an intelligence test should be to insure reliability is a question that has been frequently raised. The use of intelligence tests in industry has often been hindered by the length of time required for giving them. The Army Alpha test requires forty-five minutes, the Thorndike Comprehensive Tests approximately three hours. An investigation has been made with an intelligence test devised by the Bureau of Personnel Research of the Carnegie Institute of Technology, to determine the effect of greatly shortening the time of the test upon: (1) standing as determined by the full time test and (2) correlation with a practical criterion. The results show that an intelligence test like Personnel Bureau Test VI can be considerably shortened with practically no loss of reliability or predictive value. Accordingly an adaptation of the test to the time available—whether five minutes or a half hour—has been secured.

Personnel Bureau Test VI is a simple modification of the Army Alpha intelligence test.[1] Six parts of the Alpha have been "scrambled" to make a continuous test, the oral directions and the common sense questions only being omitted. The remaining items were arranged in "spiral omnibus" form, i. e., they become progressively harder as the test advances. The whole series is continuous, since complete directions covering all the classes of items appear at the beginning, after which the subject works steadily without interruption for the entire allotted time. The test in this form has been given to more than a thousand salesmen and office workers and is in regular use in the employment offices of the Westinghouse Electric and Manufacturing Co., the W. T. Grant Co., the Phoenix Mutual Life Insurance Co. and others.

In determining the reliability and predictive value of short intelligence tests, experiments have been made with two groups of students in the School of Life Insurance Salesmanship at Carnegie Institute of Technology. The groups consisted almost entirely of men between the ages of twenty-one and

[1] A description of this test has been previously published. Bingham, W. V. Some Group Tests. *Psychol. Bull.*, Feb., 1920, p. 57.

forty-five, all of whom were eager to meet the requirements for success in this school.

The testing of these two groups followed, with one modification, the usual procedure of allowing four minutes for reading the directions and illustrations and twenty-five minutes for the test itself. In response to a signal given by the experimenter at intervals of five minutes each, the subjects made a check mark opposite the item upon which they were then working. The papers therefore furnished five scores, as follows:

(1) Score for the first five minutes.
(2) Score for the first ten minutes.
(3) Score for the first fifteen minutes.
(4) Score for the first twenty minutes.
(5) Total score on the complete test of twenty-five minutes.

In the first experiment Test VI Form A was given to the first group, forty-eight students, in October, 1920. The relations between partial scores and total scores are shown in the following table of correlations:

	Total Score	P.E.
Score for 5 minutes	.88	±.02
Score for 10 minutes	.92	±.02
Score for 15 minutes	.96	±.01
Score for 20 minutes	.97	±.01

In the second experiment Test VI Form B was given to the second group, eighty students, in January, 1921. The correlations between partial scores and total scores were as follows:

	Total Score	P.E.
Score for 5 minutes	.89	±.01
Score for 10 minutes	.93	±.01
Score for 15 minutes	.97	±.01
Score for 20 minutes	.98	±.01

The scatter diagrams showed that the close relationship between partial and total scores held true of all the individuals, not a single widely diverging case being found.

To test further the reliability of the shortened tests the partial scores were correlated with an objective criterion, viz., the grades obtained in all school courses taken by the students. According to the system in use at the Carnegie Institute of Technology, the school grades were translated into quality points. A grade of A was given four points for quality, a grade of B three points, a grade of C two points, a grade of D one point, grades of E and F no points for quality. As each person in the insurance school took exactly the same

courses the total quality points were directly comparable for all the students.

The correlations of partial scores with quality points earned in the school courses were as follows:

	Quality Points	P.E.
Score for 5 minutes	.53	±.07
Score for 10 minutes	.54	±.07
Score for 15 minutes	.63	±.06
Score for 20 minutes	.63	±.06
Total Score	.62	±.06

The results show that an intelligence test of only five minutes is fairly proportional to the complete twenty-five minute test, a test of ten minutes more closely proportional. By the use of the regression equations scores in the short time tests can be converted into approximate full time scores and thus be compared with established norms. In order to get the intelligence score for the complete twenty-five minute test:

Multiply the score made in *five* minutes by 2.03 and add 42; or

Multiply the score made in *ten* minutes by 1.38 and add 28; or

Multiply the score made in *fifteen* minutes by 1.15 and add 13; or

Multiply the score made in *twenty* minutes by 1.02 and add 11.

This means that the intelligence test can be made short or long according to the time at the disposal of the examiner.

For practical purposes in predicting school success, the *fifteen minute test is just as satisfactory and reliable as the longer test.* Though the reliability increases slightly with the longer periods, the fast worker and the slow one alike maintain their gait throughout the test. Differences in ability appear in the first five minutes of a test, and these differences in ability are roughly proportional throughout. To continue an intelligence test of this character beyond a point of proven reliability is merely to duplicate results already secured.

BOOK REVIEWS

Lehrbuch der experimentellen Psychologie. By J. Fröbes. S. J. Herdersche Verlagshandlung, Freiburg. Bd. I, 1916, 605 pp.; Bd. II, 1920. 704 pp.

These two volumes constitute a remarkably complete compilation of psychological investigations culled from a wide and varied domain and done with a thoroughness typically Teutonic. The author's purpose was to bring together the material of experimental psychology as it exists to-day in a manner at once comprehensible for introductory work and suitable as a point of departure for advanced specialization, somewhat after the fashion of Tigerstedt's *Lehrbuch der Physiologie.* A glance at the wealth of matter in the book shows that the author has been far more inclusive than the title might indicate, for branches of psychology other than experimental are treated at considerable length.

In the first volume Fröbes deals with the simpler psychological processes. The order of treatment is conventional. The first section, following an introductory chapter on the object and methods of empirical psychology, discusses the nature and conditions of sensation in general, specific nerve energies, correspondence between sensation and stimulus, psychophysical parallelism, and leads directly to a survey of the special senses in section two. Here we find detailed treatment of the sensory phenomena of each sense department. The chapter on vision, *e. g.*, begins with the structure of the eye and takes up in turn the psychological arrangement of visual sensations, the relation of color sensations to the stimulus, color mixture, adaptation, indirect vision, contrast, after-images, the temporal course of visual sensations, color blindness, and concludes with a statement of visual theories. No less complete, although not so lengthy, are the chapters on audition, taste and smell, tactual sensations, kinaesthetic and static sensations, and organics. A chapter on the simple feelings closes section two. Section three, which is devoted to perception, deals with simple and complex ideas (*Vorstellungen*), tonal fusion, consonance and dissonance, visual and tactual space perceptions, and perceptions of time and movement. An excellent presentation of the subject of psychophysics is given in the next section. The development and practice of psychophysics from Fechner down through Wirth and Urban—but not Thomson!—with a confessed predilection for the views of G. E. Müller, are outlined without undue sidesteps into points of controversy. Fröbes' own investigations in this department are referred to quite modestly in passing. A good chapter on correlation brings the section to an end. An hundred page section on the association of ideas completes the first volume.

The second volume leaves the field of experimental psychology for those wider domains which serve as supports for language, aesthetics, criminology, sociology, education, medicine, law, *etc.* In the first volume the author was well at home in his own special branch of study: lacking this intimate familiarity with certain of the subject matter of the second volume he had the sagacity to seek expert advice from *Fachmänner.* An opening section treats of disturbances of association and cerebral localization; the so-called higher mental processes,

attention, the self, memory, reasoning, imagination, come in for detailed discussion; over an hundred pages are devoted to the emotions, aesthetic feelings, mimicry, and physiognomy; well-known authorities are quoted at length in chapters on the will, morals, customs, religion, personality, mental development, and social psychology; and the volume is brought to a close with some 125 pages on psychopathology. A valuable survey for those wishing an orientation into the topics treated!

This work of Fröbes is unquestionably the most useful source of reference for general psychology which has appeared to date. It would seem, however, that a Handbook of Psychology would be a more fitting title than a Textbook of Experimental Psychology. It does not attempt to be systematic and to that extent would be dangerous in the hands of those beginners who have no guidance from a particular point of view. But to those students who have gotten grounded somewhere in psychology and want reliable reference to less familiar topics, this work will come as a tremendous contribution. The author has selected his literature wisely and widely (some 400 references to writings in English, 230 of which are by Americans!) and has not littered the pages with inverted commas, colons, and semicolons, which verbatim quotations necessitate, but has given in clear, concise form (sometimes painfully concise: *ebenso* Witasek) an epitome of the views of authorities cited.

July, 1921, Clark University C. C. PRATT

Fundamentals of High School Mathematics. A Textbook Designed to Follow Arithmetic. By HAROLD ORDWAY RUGG and JOHN ROSCOE CLARK, both of the Lincoln School, Teachers College, Columbia University. Cloth xv+ 368 pages. Illustrated. Price $1.80. Also Answer Book, 16 cents. Yonkers-on-Hudson, New York: World Book Company.

Every teacher or school administrator who is interested in improving the educational value of the first-year high school mathematics course should examine this new book. Unlike many recent texts in algebra this one has a real excuse for being.

The authors have assembled in this ninth grade course the most important mathematical tools and notions which all children should know. The excessive manipulation of symbolisms of formal algebra has been replaced by material which is, both from the mathematical and social point of view, of far greater value. Specifically, graphs, methods of direct and indirect measurement, the properties of the right triangle, and a comprehensive treatment of the concept of relationship between changing quantities (*i. e.* functionality) have replaced the elaborate treatment of factoring, fractions, and operations with long polynomials.

The selection of material for this course is quite in agreement with the recommendations of the National Committee on Mathematical Requirements. The authors are to be congratulated upon having sensed fully the best thought of the day in the reorganization of first-year high school mathematics courses.

Among the unique impressions made upon the reviewer, the following stand out most prominently:

(1) The careful explanations and development of new processes;
(2) The wholesome omission of formal material;
(3) The excellent presentation of word-problems;

(4) The unique organization of special products and factoring;
(5) The "timed practice-exercises" for developing skill in essential tool processes; and
(6) The emphasis upon the notion of relationship between variable quantities.

Manhood of Humanity; The Science and Art of Human Engineering. By ALFRED KORZYBSKI. E. P. Dutton & Co., New York, 1921, p. 264.

This is a book of very wide scope embracing all human problems. It is a study of man and the problems of humanity from a scientific, especially a mathematical, point of view. The author makes a plea for the establishment of a "Science and Art of Human Engineering" dealing with human problems not by means of a study of metaphysical essences, but of actual facts, and reduces these facts to mathematical laws. The author attempts to explain social problems by an application of mathematical reasoning and mathematical law. "Engineering is the co-ordinated sum-total of human knowledge gathered through the ages, with mathematics as its chief instrument and guide."

In his first chapter, the author deals with a new concept of life and makes use of a very interesting mathematical analogy which although in the opinion of the reviewer, somewhat overdrawn, nevertheless serves to bring out in a very striking manner how the progress of the social sciences has lagged behind that made by the natural sciences. The second chapter takes up The Childhood of Humanity and shows how metaphysical speculation, selfish political philosophies and sectarian opinions have hindered the progress of the social sciences and thus prolonged the period of humanity's childhood right up to the present day. In the next chapter the author gives an unique classification of the types of life which points out vividly the relation of man to other classes of life and shows how the great importance of man lies in his "time-binding capacity." This classification is fundamental in the development of the author's theory. The entire next chapter is devoted to a systematic inquiry into the nature of man remarkably free from metaphysics and dogma in an attempt to arrive at a scientific understanding of the function and sphere of man. This is perhaps the most important chapter in this important book and the concept here developed by the author is of far-reaching importance in the scientific solution of humanity's problems.

Chapter V Wealth and Chapter VI Capitalistic Era outline the method of approach of the new science of Human Engineering to these fundamental exonomic problems. The author points out that economic ideas "however false and harmful, are protected alike by habit and by the inborn conservatism of many minds" and how these old ideas hinder the natural growth of our social sciences. These economic problems are attacked in a clear, logical way and an attempt is made to get at fundamental economic concepts, reducing them to mathematical law. This is a practical application of the author's theory to the field of economics. Chapter VII deals with the Survival of the Fittest in a useful way and points out the difference in the way survival works in animals and in men. Chapter VIII, The Elements of Power, endeavors to explain power by mathematical and mechanical methods and takes up the German philosophy of power that had so much to do in bringing on the World War. In Chapter

IX, Manhood of Humanity, the author shows that a scientific basis for the study of the social sciences will initiate the Manhood of Humanity, when the social sciences will prevent the energies of man from being wasted by ignorance and selfishness, by competitive conflict characteristic of beasts, but will show that these energies are more than sufficient to produce a high order of increasing prosperity everywhere throughout the world. The author gives a necessarily vague but suggestive outline of a plan for science and art of Human Engineering to guide the large affairs of the new state of humanity's manhood. This state is to be controlled by scientists and not by petty politicians. "The world will have uninterrupted, peaceful progress when and only when the so-called social 'sciences'—the life-regulating 'sciences' of ethics, law, philosophy, economics, religion, politics, and government are technologized; when and only when they are made genuinely scientific in spirit and method."

The book also contains Appendix I which takes up Mathematics and Time-Binding; Appendix II, Biology and Time-Binding; and Appendix III, Engineering and Time-Binding giving a choice bibliography for each appendix.

Manhood of Humanity is a valuable attempt to a true conception of what human beings really are—an attempt to initiate the scientific study of man—a science and art that will know how best to direct the energies of man to the advancement of human welfare. It is written in a clear, logical, stimulating style and the theory advanced by the author will probably arouse much controversy. It is an attack of a very old subject from a new point of view and is a truly remarkable contribution toward the scientific study of humanity and should command the attention of all interested in humanity's problems.

MAX MEENES

Methods and Results of Testing School Children. By EVELYN DEWEY, EMILY CHILD and BEARDSLEY RUML. E. P. Dutton & Co., New York, 1920, p. 176.

This book is the result of a Psychological Survey conducted in the New York City Public Schools in February, 1915. "The immediate purpose of the investigation was to obtain norms for a series of tests for New York public school children in the poorer and more congested portions of the city, and a basis for further study of the value of mental tests for improved school procedure. It was hoped that the results would justify the establishment of a school clinic for normal children to assist the teacher in meeting the problems of individual pupils." Accordingly, a series of tests was devised to measure numerically five differences of performances between children of the same chronological age and it was aimed to accomplish this by presenting situations requiring mental adjustments of the type that a child uses in adapting himself to his environment. The authors confess that although improvements could have been made at every stage of the investigation, they limited the changes almost wholly to the elimination of tests that were unsatisfactory in some phase of administration in order to avoid an impracticable repetition of the tests. It is perhaps significant to note that in giving the test the authors adopted a method in which the maximum amount of rigidity and uniformity were prescribed, under the conviction that a really quantitative measure of the child's adjustment was their goal and would justify the limits in flexibility which this method would imply.

The authors based their investigation on results obtained from Jewish subjects and aimed further to collect a parallel set of results in schools where children of other nationalities predominated. Pupils between the ages of eight and thirteen were tested. A large variety of "tests" were given to the children individually and under standard conditions. An effort was made to avoid the child's receiving the impression that he was being classified according to his performance in the tests or that he was undergoing a school examination. The testing was never begun until the child felt at home with the tester and interested in doing as he was asked. His mistakes were never corrected, the tester maintaining an encouraging and satisfied manner no matter what the response." This was not a group intelligence test, but an individual performance scale containing a number of separate special tests such as cancellation, tapping, tests of manual dexterity, cart construction and some tests from the Yerkes-Bridges and from the Binet scales. The authors give a description of these various tests and the directions for giving them, together with the methods used for working up the results. The results are treated statistically in the many tables and charts contained in the book. An added feature is a number of charts showing sex differences brought out by the tests used in this investigation. A home investigation of the children tested was also made to provide data for an extensive social study of the homes from which the children came and the social data thus obtained is contained in tabular form. To complete the survey, all the children tested were given a thorough physical examination and this book contains tables of norms obtained in the physical examination of these children.

The investigation is rather thorough as far as it goes and the work was necessarily slow and painstaking, but such an investigation must be pushed further to include larger numbers of children of all nationalities and all environmental conditions so that satisfactory norms may be established and Psychological Clinics formed in connection with the school departments of various cities. MAX MEENES.

NOTES

INDUSTRIAL RESEARCH LABORATORIES

Research facilities and the development activities of American industries are to be described in the forthcoming revision of *Bulletin of the National Research Council* Number 2, "Research laboratories in industrial establishments of the United States of America." Only 300 such laboratories were listed in the first edition but it is hoped that several hundred new names will appear in the revision and that a more nearly complete reference list will thus become available. The general demand for the first edition of the Bulletin shows the wide interest in this subject, and the importance of having every laboratory which devotes even a portion of its time to research properly listed.

The Council requests information from directors of research who have not already supplied it. The following data are wanted: Name and address of firm and address of laboratory; name of director of research; number on laboratory staff (classified as chemists, engineers, bacteriologists, etc.); approximate proportion of time spent on research; chief lines of research; unusual features of equipment; research laboratory space; date of organization of research laboratory and annual expenditure for research. Confidential information is not desired.

It is also requested that librarians in the service of the industries please bring this notice to the attention of the proper officials in their organizations.

This material should be furnished as promptly as possible to the Research Information Service, National Research Council, 1701 Massachusetts Avenue, Washington, D. C.

REPORT OF PSYCHOLOGICAL WORK IN THE UNITED STATES ARMY

A complete report of the history, methods and results of psychological examining in the United States Army has been recently published in the Memoirs of the National Academy of Sciences, Volume 15, 1921. The report is edited by Lieut. Col. Robert M. Yerkes, Chief of the Division of Psychology, as an official document for the Surgeon General of the Army. It consists of three parts bound in a single volume. Part I, presenting the official history of the development of the service and its activities during the war, is supplemented by reproductions of all of the printed materials devised and used in conducting psychological examinations. Part II includes a complete account of the preparation of methods, their characteristics, and their evaluations as practical procedures. In Part III the results of examining are summarized.

The entire report may be obtained from the Superintendent of Documents, Government Printing Office, Washington, D. C., at $1.75 per copy. It appears in quarto size under the title "Psychological Examining in the United States Army," and includes VI+890 pages.

The following books and pamphlets have been received:

Charles S. Myers. Mind and Work. G. P. Putnam's Sons. New York and London. 1921, pp. 175.

Stewart Paton. Human Behavior. Charles Scribner's Sons. New York. 1921. pp. 465.

Rugg and Clark. Fundamentals of High School Mathematics. World Book Co. Yonkers-on-Hudson, N. Y. 1921. pp. 368.

Harl Roy Douglass. The Derivation and Standardization of a Series of Diagnostic Tests for the Fundamentals of First Year Algebra. University of Oregon Publication, Eugene, Ore. Vol. 1, No. 3. April, 1921.

Arnold Gesell. Exceptional Children and Public School Policy. Yale University Press, New Haven, Conn. 1921.

Clara H. Town. Analytic Study of a Group of Five and Six-Year-Old Children. University of Iowa Studies, Iowa City, Ia. Vol. I, No. 4. May 1, 1921.

Bird T. Baldwin. The Physical Growth of Children from Birth to Maturity. University of Iowa Studies. Vol. I, No. 1. June 1, 1921.

Bassett and Porteus. Sex Differences in Porteus Maze Test Performance. Reprint from the Training School Bulletin. Vineland, N. J. No. 22. Dec., 1920.

Bureau of Juvenile Research. Review of the Work, 1918-1920. Columbus, Ohio. Publication No. 19. Feb., 1921.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Cornell Univ. Subscription $6.00. 600 pages annually. Edited by Edward Bradford Titchener. Quarterly. General and experimental psychology. Founded 1887.

Pedagogical Seminary—Worcester, Mass.: Ritter McLoughlin. Subscription $5. 575 pages annually. Edited by G. Stanley Hall. Quarterly. Pedagogy and educational psychology. Founded 1891.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $4.25. 480 pages annually. Edited by Howard C. Warren. Bi-monthly. General. Founded 1894.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $3. 720 pages annually. Edited by Shepherd I. Franz. Monthly. Psychological literature. Founded 1904.

Psychological Monographs—Princeton, N. J.: Psychological Review Company. Subscription $5. [illegible] vol. 500 pp. Ed. by [illegible]. Published without fixed dates, each issue one or more [illegible]. Founded 1895.

Psychological Index—Princeton, N. J.: Psychological Review Company. Subscription $1.50. 200 pp. Edited by Madison Bentley. An annual bibliography of psychological literature. Founded 1895.

Journal of Philosophy, Psychology, and Scientific Methods—New York: Sub-station 84, N. Y. Subscription $4. 728 pages annually. Edited by R. J.E. Woodbridge and Wendell T. Bush. Bi-weekly. Founded 1904.

Archives of Psychology—Sub-station 84, N. Y. [illegible] Psychology. Subscription $5. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal Psychology—Boston: Richard G. Badger. Subscription $5. 480 pages annually. Edited by Morton Prince. Bi-monthly. Journal of abnormal psychology. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $2.50. 288 pages ann. Ed. by Lightner Witmer. Without fixed dates (9 numbers). Orthogenics, psychology, hygiene. Founded 1907.

Training School Bulletin—Vineland, N. J.: The Training School. Subscription $1. 180 pages ann. Ed. by [illegible] Johnstone. Monthly (10 numbers). Psychology and training of defectives. Founded 1904.

Journal of Educational Psychology—Baltimore: Warwick & York. Subscription $[illegible] annually. Monthly (9 numbers). Managing Editor, Harold O. [illegible] (Educational Psychology Monographs. Published separately at varying prices. [illegible]) Founded 1910.

The Behavior Monographs—Cambridge, Mass.: Harvard Univ. Subscription $5. [illegible] pages per volume. Edited by John B. Watson. Published [illegible] dates, each number a single research.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W. Subscription $6. 500 pages annually. Edited by W. A. White and S. E. Jelliffe. Psychoanalysis. Quarterly. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J.: Psychological Review Company. Subscription $4.25. 480 pages annually. Ed. by John B. Watson. Bi-monthly. Experimental. Founded 1916.

Journal of Applied Psychology—Worcester, Mass.: Florence Chandler. Subscription $5. 400 pages annually. Edited by James P. Porter and William D. Scott. Quarterly. Founded 1917.

The Journal of Comparative Psychology—Baltimore: Williams & Wilkins Co. Subscription $5. 400 pages, approximately, per volume. Edited by Knight Dunlap and Robert M. Yerkes. Bi-monthly. Founded 1921.

Journal of International Relations—Worcester, Mass.: Louis N. Wilson. Subscription $3. 300 pages annually. Edited by George H. Blakeslee and [illegible] Barnes, assisted by a Board of some [illegible] leading editors. Quarterly. Founded 1910.

Penn State University Pattee Library
Penn State University
University Park, PA